

Align seg 1/1 to: ABA47096 from: 1 to: 120

1 GCCAGATTCAATCCCATGTCGAGGCTTCGCCGCAAAATTACAGTGG 50
|||||
1 AlaArgpheAsnProIle**trpArgLeuArgAlaGlnIleAsnSerG1 17
51 GGAGTCNTCCCAACAAAGTACATGCTGACGTGAGACAGAGTTGGG 100
|||||
17 yJusSerProPheLeuProAsnMetIleuGlnIleuGlnIleuSerTTPG 34
101 GTGCAATCCCAACCAAGTCCAGAGTCCAGAACGATGATTTTAATACC 150
|||||
34 lYArgGlnSerThrGlnValGlnGlnSerIleAspGlyIlePheAsnThr 50
|||||
151 CAATTAAGATTGACAGATTTCGCCGCGTAACCTTGTGACGNTGACCAATGT 200
|||||
51 GlnIleArgLeuGlnIleSerAlaGlnIleAsnPheValThr**SerAsnVa 67
201 TCGCAGCGTATCTCAGTCAGTCAGTCATGTTGTGCAATGCAATGATGC 250
|||||
67 lArgAspValIleSerSerLeuAlaIleMetLeuPheGlnCysSerG1YA 84
251 GGCCATTCTCCTCTCTCGACACCCCTTCGCCGCTGCTCTCAAGTCCTGC 300
|||||
84 rGProPheSerSerLeuAsnProIlePheTrpSerProLeuLeuArgSerVal 100
301 GTGCAATGCGCGCAAGCAATGTCACCTGCACTGTTTCCCAACCCAGCCGCG 350
|||||
101 ValAspAlaIleAsnAspValThrCysThr**SerGlnProThrValAr 117
351 CATCGTA 357
|||||
117 g1IleVal 119

seq_name: /stid1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:ABA47097
seq_documentation_block:
ID ABA47097 standard; Protein: 174 AA.
XX ABA47097;
AC
XX 16-MAY-2001 (first entry)
XX
XX Korean mistletoe lectin #2.
XX
XX Isoform: A-chain; B-chain: biosynthesis; lectin; Korean mistletoe;
XX KML; tumour: KM-110; KML-C; KMBP; KML-IU; KML-IIU;
XX heparin binding protein.
XX
XX Viscum album coloratum.
XX
XX EPI074560-A2.
XX
XX 07-FEB-2001.
XX
XX 27-JUL-2000; 2000EP-0402168.
XX
XX 27-JUL-1999; 99KR-0030638.
XX
XX (MIST-1) MISTLE BIOTECH CO LTD.
XX
XX Kim J, Song S, Suh B, Lee K, Doo M, Kwak J, Song B, Yoon T;
XX Kang T, Park C;
XX WPI: 2001-171044/18.
XX N-PSDB: AAC85479.
XX
XX Novel lectin proteins isolated from Korean mistletoe, useful for
XX enhancing immunity and effectuating anti-tumoral activity -
XX
XX Claim 32: Page 34: 62pp; English.
XX
XX The sequences given in ABA47096-97 are lectins isolated from Korean

CC Mistletoe. Korean mistletoe lectins (KML) are useful for enhancing
CC immunity and for treating tumours. The KML's are isolated from a
CC protein fraction derived from the leaves, stems and fruits of Korean
CC mistletoe, which is designated KM-110. One of the isolates, KML-C was
CC shown to be effective against colon 26-M3.1 carcinoma and L5178Y-M25
CC lymphoma.
XX
XX Sequence 174 AA:
50

alignment_scores:
Quality: 408.50 Length: 122
Ratio: 4.045
Percent Similarity: 82.787 Percent Identity: 69.672

alignment_block:
US-09-627-165b-13 x ABA47097 ..

Align seg 1/1 to: ABA47097 from: 1 to: 174

1 GCCAGATTCAATCCCATGTCGAGGCTTCGCCGCAAAATTACAGTGG 50
|||||
53 AlaArgpheAsnProIlePheTrpArgAlaArgGlnIleAsnSerG1 69
51 GGAGTCNTCCCAACAAAGTACATGCTGACGTGAGACAGAGTTGGG 100
|||||
69 yJusSerProPheLeuProAsnMetIleuGlnIleuGlnIleuSerTTPG 86
101 GTGCAATGCGCGCAAGCAATGTCACCTGCACTGTTTCCCAACCCAGCCGCG 150
|||||
86 lYArgGlnSerThrGlnValGlnGlnSerIleAspGlyIlePheAsnThr 102
151 CAATTAAGATTGACAGATTTCGCCGCGTAACCTTGTGACGNTGACCAATGT 200
|||||
103 ProPheArgLeuGlnIleSerThrGlnIleAsnPheValThrIleAsnSerVa 119
201 TCGCAGCGTATCTCAGTCAGTCAGTCATGTTGTGCAATGCAATGATGC 250
|||||
119 lArgAspValIleAsnSerLeuAlaIleMetLeuPheValCysArgAspA 136
251 GGCCATTCTCCTCTCTCGACACCCCTTCGCCGCTGCTCTCAAGTCCTGC 300
|||||
136 rGProPheSerSerLeuAsnProIlePheTrpSerProLeuLeuArgSerVal 152
301 GTGCAATGCGCGCAAGCAATGTCACCTGCACTGTTTCCCAACCCAGCC 341
153 LeuGlnAsnSerGlyAlaValAspAspValThrCysThrAlaSerGluPr 169
342 CACCGTGGCAGTCGTA 357
|||||
169 oThrValArgIleVal 174

seq_name: /stid1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:AAW10021
seq_documentation_block:
ID AAW10021 standard; Protein: 564 AA.
XX AAW10021;
AC
XX 18-DEC-1997 (first entry)
XX
XX Prepro mistletoe lectin.
XX
XX Mistletoe; lectin; cytotoxic; A chain; B chain; dimer.
XX
XX Viscum album.
XX
XX EP751221-A1.
XX
XX 02-JAN-1997.
XX
XX 26-JUN-1995, 95EP-0109949
XX

```
PR 26-JUN-1995; 95EP-0109949.
XX (MADU ) MADNUS KOELN AG.
XX Baur A, Eck J, Lentzen H, Zinke H;
XX WPI: 1997-054678/06.
XX DR N-PSDB: AAT70473.
XX PT Nucleic acid encoding pre:pro form of mistletoe lectin - for
XX therapeutic or diagnostic use
XX PS Claim 12; Fig 4c; 30pp; German.
XX CC Mistletoe lectin is a cytotoxic agent that has been used for tumour
XX therapy. It can be used in immunotoxins and medicaments. Nucleic
XX acid fragments can be used in diagnostic methods. Mistletoe lectin (
XX AAT70473) comprises an A chain (AAT70474) and a B chain (AAT70475).
XX SQ Sequence 564 AA;

alignment_scores:
    Quality: 387.00 Length: 119
    Ratio: 3.990 Gaps: 1
    Percent Similarity: 81.513 Percent Identity: 67.227

alignment_block:
US-09-627-165B-13 x AAM10021 ..

Align seg 1/1 to: AAM10021 from: 1 to: 564

1 GCCAATTCATCCCATCCTGTCGAGCGCTTCCGCGCAATTAACAGTGG 50
|||||
200 AlaArgPheAsnProIleLeuTrpArgAlaArgGlnTyrIleAsnSergl 216
51 GGAGCNCNTCCACCAACATGTACATGCTGCAGCTGCAGAGCGAGTGGG 100
|||||
216 ValAserPheLeuProAspValTyrMetLeuGlnLeuGlnTyrSertIrgp 233
101 GTGCAACATCCAGCCCAAGTCAGACGTCAGAGATGGCATTTTAATACC 150
|||||
233 TygInGlnSerThrGlnValGlnIleSertThrAspGlyValPheAsnAsn 249
151 CAAATAGATTGCAGATTTCCGCCGGTAACCTTGTGAGAGNTGACCAATGT 200
|||||
250 ProIleArgLeuAlaIleProProGlyAsnPheValThrLeuThrAsnVa 266
201 TCGGAGCGTATTCGACGCTTGGCGATCATGTTGTTCGAATGCGAGTGTC 250
|||||
266 lArgAspValIleAlaSerLeuAlaIleMetLeuPheValIcysGlyGln 283
251 GGCATTTCTCTCTGACACCCCTTCCGCGCTGCTCCTAAGSGTCGTC 300
|||||
283 rProIleSertSergIleValAlaIrgTyrTrpProLeuValIleIArgProVal 299
301 GTGATGCGGCCAAGATGTACATGCTGCACTNTTTCGAAACCACGCTGCG 350
|||||
300 Ile.....AlaAspAspValThrCysSerAlaSerGlnProThrValAr 314
351 CATCGTA 357
314 gIleVal 316

seq_name: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:AAW90127
seq_documentation_block:
ID AAW90127 standard; Protein; 564 AA.
XX
XX AAW90127;
XX
XX 30-APR-1999 (first entry)
XX
```

```
DE Mistletoe lectin prepro-protein.
XX
XX MW: mistletoe; lectin; ML: transgenic plant; glycosylation;
XX dimer; immunotoxin; large-scale production; diagnosis; therapeutic;
XX cancer.
XX Viscum album.
XX EP84388-A1.
XX PD 16-DEC-1998.
XX
XX 26-JUN-1995; 95EP-0109949.
XX
XX 26-JUN-1995; 95EP-0109949.
XX PR 26-JUN-1995; 98EP-0105660.
XX
XX (MADU ) MADNUS KOELN AG.
XX Baur A, Eck J, Lentzen H, Zinke H;
XX WPI: 1999-026582/03.
XX DR N-PSDB: AAV74182.
XX
XX New transgenic plant expressing mistletoe lectin - useful for
XX producing recombinant lectin in e.g. cancer diagnosis and therapy
XX
XX Claim 1a; Fig 4c; 30pp; German.
XX
XX This invention describes a novel transgenic plant transformed with a
XX vector capable of encoding a mistletoe (Viscum album) lectin
XX preprotein or a biologically active fragment. The specification
XX also describes a polypeptide produced by a plant where the polypeptide
XX exhibits at least one enzymatic modification other than the glycosylation
XX that occurs in Viscum album or the polypeptide is a fusion protein, a
XX mistletoe lectin polypeptide dimer and an immunotoxin comprising the
XX polypeptide or the polypeptide dimer. The plants are used for large-scale
XX production of mistletoe lectin for diagnostic or therapeutic purposes
XX (e.g. in cancer therapy). This sequence represents the mistletoe lectin
XX used in the method of the invention.
XX SQ Sequence 564 AA;

alignment_scores:
    Quality: 387.00 Length: 119
    Ratio: 3.990 Gaps: 1
    Percent Similarity: 81.513 Percent Identity: 67.227

alignment_block:
US-09-627-165B-13 x AAW90127 ..

Align seg 1/1 to: AAW90127 from: 1 to: 564

1 GCCAATTCATCCCATCCTGTCGAGCGCTTCCGCGCAATTAACAGTGG 50
|||||
200 AlaArgPheAsnProIleLeuTrpArgAlaArgGlnTyrIleAsnSergl 216
51 GGAGCNCNTCCACCAACATGTACATGCTGCAGCTGCAGAGCGAGTGGG 100
|||||
216 ValAserPheLeuProAspValTyrMetLeuGlnLeuGlnTyrSertIrgp 233
101 GTGCAACATCCAGCCCAAGTCAGACGTCAGAGATGGCATTTTAATACC 150
|||||
233 TygInGlnSerThrGlnValGlnIleSertThrAspGlyValPheAsnAsn 249
151 CAAATAGATTGCAGATTTCCGCCGGTAACCTTGTGAGAGNTGACCAATGT 200
|||||
250 ProIleArgLeuAlaIleProProGlyAsnPheValThrLeuThrAsnVa 266
201 TCGGAGCGTATTCGACGCTTGGCGATCATGTTGTTCGAATGCGAGTGTC 250
|||||
266 lArgAspValIleAlaSerLeuAlaIleMetLeuPheValIcysGlyGln 283
```


PS Disclosure: Fig 4B: 78pp: German.

XX This invention describes a novel mistletoe lectin (I) and its fragments
CC which have antitumour and immunostimulatory activity. The A-chain (MLA)
CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of
CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and
CC lymphokine-producing macrophages, so stimulate immunity (I) and its
CC fragments are used to treat uncontrolled cell growth (particularly
CC cancers) and if they lack cytotoxicity, to increase the strength of the
CC immune response, particularly to a co-administered antigen
CC (tumour-associated, bacterial or viral). The method allows production of
CC mistletoe lectin, and its individual chains, in many different isoforms
CC and on a large scale, at any time of the year. Recombinant products are
CC free from toxins present in natural mistletoe extracts. This sequence
CC represents a fragment of a mistletoe lectin I protein variant.

XX Sequence 532 AA;

Alignment_scores:

Quality: 386.00 Length: 119
Ratio: 3.979 Gaps: 1
Percent Similarity: 81.513 Percent Identity: 67.227

Alignment_block:

US-09-627-165B-13 x AAY25982 ..

Align seg 1/1 to: AAY25982 from: 1 to: 532

1 GCCGATTCATCCATCMTGTGAGGCTTCGCCGCAATTAACAGTGG 50
|||||
167 AlalrghphasnProilleutrparglAlarglIntyrileasnsergl 183
51 GGAATCMTTCACCAACATGATACATCTGACGTGAGCAGATGGG 100
1 ||| |||
183 yAlSerPheLeuProAspValTyrMetLeuGlulLeuGlutThrSertips 200
101 GTCGACAAATCCAGCAGTCCAGCAGTCCAGATGGATTTATATAC 150
|||||
200 lyGlnGlnSerThrGlnValGlnHisSerThrAspValAlaPheAsn 216
151 CAATTAAGATTGCAGATTTCGCCGATCTTGTGACGNTGCAATGT 200
|||||
217 ProileArGleuAlaIleProProGlyAsnPhaValThrLeuThrAsnVa 233
201 TCGGACGTGATCTCCAGCTTCGCGATCATGTGTTGCAATGCAATGTC 250
|||||
233 lArGAspValIleAlaSerLeuAlaIleMetLeuPhaValCylGlyUA 250
251 GGCATTCCTCCTCTCGACACACCTTCGCCGCTGCTGCTTAAGTCCGTC 300
|||||
250 rGProSerSerSerAspValArGlyTyrProLeuValIleArGProVal 266
301 GTGATGCGCGCAGACATGTCACTGCACTTTTCCAGACCCACCGTCG 350
:::
267 Ile.....AlaAspAspValThrCysSerAlaSerIleProThrValAr 281
351 CATCGTA 357
|||||
281 gIleVal 283

seq_name: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT: AAY25970

seq_documentation_block:

ID AAY25970 standard; protein: 533 AA.

XX AAY25970:

XX 18-OCT-1999 (first entry)

XX Mistletoe lectin protein consensus sequence 1.

XX Mistletoe; lectin; antitumour; immunostimulant; A-chain; MLA; immunity;

KW rIbozyme 28S subunit; non-cytotoxic; T-cell activation; immune response;
KM lymphokine-producing macrophage; uncontrolled cell growth; treatment;
KW cancer; cytotoxicity; antigen; isoform.

XX Viscum album.

XX Key location/Qualifiers
FH Misc-difference 15
FT Misc-difference /label= Asp, Glu
FT Misc-difference 63
FT Misc-difference /label= Gly, Gln
FT Misc-difference 66
FT Misc-difference /label= Ile, Val
FT Misc-difference 75
FT Misc-difference /label= Leu, Ala
FT Misc-difference 107
FT Misc-difference /label= Asp, Arg or none
FT Misc-difference 113
FT Misc-difference /label= Asn, Thr
FT Misc-difference 117
FT Misc-difference /label= Pro, Thr
FT Misc-difference 134
FT Misc-difference /label= Asp, Glu
FT Misc-difference 141
FT Misc-difference /label= Ser, Thr
FT Misc-difference 145
FT Misc-difference /label= Phe, Tyr
FT Misc-difference 152
FT Misc-difference /label= Thr, Ala
FT Misc-difference 177
FT Misc-difference /label= Ala, Tyr
FT Misc-difference 180
FT Misc-difference /label= Tyr, Asp
FT Misc-difference 185
FT Misc-difference /label= Ala, Glu
FT Misc-difference 191
FT Misc-difference /label= Val, Met
FT Misc-difference 219
FT Misc-difference /label= Ile, Phe
FT Misc-difference 224
FT Misc-difference /label= Pro, Ser
FT Misc-difference 225
FT Misc-difference /label= Pro, Thr
FT Misc-difference 232
FT Misc-difference /label= Thr, Ser
FT Misc-difference 236
FT Misc-difference /label= Asp, Ser
FT Misc-difference 287
FT Misc-difference /label= Asn, Ser
FT Misc-difference 290
FT Misc-difference /label= Cys, Arg
FT Misc-difference 325
FT Misc-difference /label= Gly, Asn
FT Misc-difference 364
FT Misc-difference /label= Gly, Asp
FT Misc-difference 426
FT Misc-difference /label= Gly, Gln
FT Misc-difference 435
FT Misc-difference /label= Val, Asp
FT Misc-difference 439
FT Misc-difference /label= Gln, Lys
FT Misc-difference 442
FT Misc-difference /label= Gly or none
FT Misc-difference 443
FT Misc-difference /label= Arg, Lys
FT Misc-difference 464
FT Misc-difference /label= Cys, Ser, Val
FT Misc-difference 480
FT Misc-difference /label= Ala, Gly
FT Misc-difference 481
FT Misc-difference /label= Gly, Ala
FT Misc-difference 483
FT Misc-difference /label= Ser, Gly

```

FT MISC-difference 484 /label= gly, Ser
FT MISC-difference 493 /label= Gly, Tyr
FT MISC-difference 500 /label= Gly, Tyr
FT MISC-difference 501 /label= Asn, Ser, Thr, Lys
FT MISC-difference 501 /label= Ser, Gly
FT MISC-difference 502 /label= Leu, Pro
FT MISC-difference 503 /label= Ala, Met
FT MISC-difference 504 /label= Met, Val
FT MISC-difference 533 /label= Pro, Phe
FT DE19804210-A1.
XX 12-AUG-1999.
XX 03-FEB-1998; 98DE-1004210.
XX 03-FEB-1998; 98DE-1004210.
XX (BIOS-) BIOSYN ARZNEIMITTEL GMBH.
XX Morris P, Stiefel T, Voelter W, Welters P;
XX WPI; 1999-44535/38.
XX Preparation of mistletoe lectins in heterologous systems,
XX particularly for use as anticancer agents and immunostimulants
XX
XX Claim 1, Page 25-26; 78pp; German.
XX
XX This invention describes a novel mistletoe lectin (I) and its fragments
XX which have antitumour and immunostimulatory activity. The A-chain (MLA)
XX of the mistletoe lectin binds to, and inactivates, the 28S subunit of
XX ribosomes. Non-cytotoxic forms of (I) activate T-cell and
XX lymphokine-producing macrophages, so stimulate immunity. (I) and its
XX fragments are used to treat uncontrolled cell growth (particularly
XX cancers) and if they lack cytotoxicity, to increase the strength of the
XX immune response, particularly to a co-administered antigen
XX (tumour-associated, bacterial or viral). The method allows production of
XX mistletoe lectin, and its individual chains, in many different isoforms
XX and on a large scale, at any time of the year. Recombinant products are
XX free from toxins present in natural mistletoe extracts. This sequence
XX represents a consensus sequence of the mistletoe lectin described in the
XX specification.
XX
XX SO Sequence 533 AA;

alignment_scores:      Length: 119
Quality: 373.00      Gaps: 1
Ratio: 3.885
Percent Similarity: 80.672 Percent Identity: 65.546

alignment_block:
US-09-627-165B-13 x AAY25970 ..

Align seg 1/1 to: AAY25970 from: 1 to: 533

1 GCCAGATTCATCCATCCTGTCAGAGCTTCGCCGCAATTAAGACGTGG 50
|||||
168 AAlArghshsAnProilleuTrpArg**Argln**llehsnSercl 184

51 GGAGTCNTTCACCAACATGTACAGTGCAGTGCAGACGAGTGGG 100
|||||
184 y***SerPheleuProasp***TyrMetleuGluIeuGluThrSerTrpG 201

101 GTCGCAATCCACCAAGTCAGAGTCAGTCAGATGCAATTTTAATACC 150

```

```

|||||
201 lyGlnInsSerThrGlnValGlnHisSerThrAspGluValAlaPhehsn 217
151 CAATTAAGATTCGACATTCGCCGCGTAACCTTTGTGACGNTGACCAATGT 200
|||||
218 Pro**ArghleuAlalle***GlysnPheValThrleu***AsnVa 234
201 TCGGACGTGATCTCCAGCTTCGCGATCATGTTGTTGCAATGCAAGTGTG 250
|||||
224 lArg**ValilleAlaSerleuAlalleMetleuPheValIcysGlyGlnA 251
251 GGCCATTCCTCTCTCGACACCCCTTCGCCGCTGCTCCTAAGTCCGTC 300
|||||
251 rProSerSerSerAspValArgTyrTrpProleuValilleArgProVal 267
301 GTCGATCCGCGCCCAAGCATGTGACCTGCACATTTTTCGAAACCCAGCGTC 350
|||
268 lle.....AlaAspAspValThrCysSerAlaSerGluProThrValAr 282
351 CATCGTA 357
|||||
282 gileVal 284

seq_name: /SIDSI/gcgcdata/hold_geneseg/genesegp-emb1/AA1999.DAT.AAY25973
seq_documentation_block:
ID AAY25973 standard; protein; 533 AA.
XX AAY25973;
XX
XX 18-OCT-1999 (first entry)
XX
XX Mistletoe lectin protein consensus sequence 2.
XX
XX Mistletoe lectin: antitumour; immunostimulant; A-chain; MLA; immunity;
XX lysozyme 28S subunit; non-cytotoxic; T-cell activation; immune response;
XX lymphokine-producing macrophage; uncontrolled cell growth; treatment;
XX cancer; cytotoxicity; antigen; isozyme.
XX
XX Viscum album.
XX
XX Key Location/Qualifiers
XX MISC-difference 15 /label= Asp, Glu
XX MISC-difference 63 /label= Gly, Gln
XX MISC-difference 66 /label= Ile, Val
XX MISC-difference 75 /label= Leu, Ala
XX MISC-difference 107 /label= Asp, Arg, none
XX MISC-difference 113 /label= Asn, Thr
XX MISC-difference 117 /label= Pro, Thr
XX MISC-difference 134 /label= Asp, Glu
XX MISC-difference 141 /label= Ser, Thr
XX MISC-difference 145 /label= Phe, Tyr
XX MISC-difference 152 /label= Thr, Ala
XX MISC-difference 177 /label= Ala, Tyr
XX MISC-difference 180 /label= Tyr, Asp
XX MISC-difference 185 /label= Ala, Glu
XX MISC-difference 191 /label= Val, Met
XX MISC-difference 219

```

CC		This invention describes a novel mistletoe lectin (I) and its fragments which have antitumour and immunostimulatory activity. The A-chain (MRA)
XX	DE19804210-A1.	
XX	12-AUG-1999.	
XX	03-FEB-1998;	98DE-1004210.
XX	03-FEB-1998;	98DE-1004210.
XX	(BIOSYN ARZNEIMITTEL GMBH.	
XX	Morris P, Stiefel T, Voelter W, Welters P,	
XX	WPI: 1999-445335/38.	
XX	Preparation of mistletoe lectins in heterologous systems,	
XX	particularly for use as anticancer agents and immunostimulants	
XX	Claim 4; Page 28-29; 78pp: German.	
XX		
FT	/label= Ile, Phe	
FT	224	
FT	/label= Phe, Ser	
FT	225	
FT	/label= Pro, Thr	
FT	232	
FT	/label= Thr, Ser	
FT	236	
FT	/label= Asp, Ser	
FT	287	
FT	/label= Asn, Ser	
FT	290	
FT	/label= Cys, Arg	
FT	323	
FT	/label= Gly, Asn	
FT	364	
FT	/label= Gly, Asp	
FT	427	
FT	/label= Gly, Gln	
FT	435	
FT	/label= Val, Asp	
FT	439	
FT	/label= Gln, Lys	
FT	442	
FT	/label= Gly, none	
FT	443	
FT	/label= Arg, Lys	
FT	464	
FT	/label= Cys, Ser, Val	
FT	480	
FT	/label= Ala, Gly	
FT	481	
FT	/label= Gly, Ala	
FT	483	
FT	/label= Ser, Gly	
FT	484	
FT	/label= Gly, Ser	
FT	493	
FT	/label= Gly, Tyr	
FT	500	
FT	/label= Asn, Ser, Thr, Lys	
FT	501	
FT	/label= Ser, Gly	
FT	502	
FT	/label= Ieu, Pro	
FT	503	
FT	/label= Ala, Met	
FT	504	
FT	/label= Met, Val	
FT	533	
FT	/label= Pro, Phe	
PN		
PD		
PF		
PR		
PA		
PI		
PX		
PY		
RN		
RP		
RT		
RV		
SF		
SN		
SP		
SR		
SS		
ST		
SV		
SW		
TX		
TY		
UZ		

CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of
CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and its
CC lymphokine-producing macrophages, so stimulate immunity. (I) and its
CC fragments are used to treat uncontrolled cell growth (particularly
CC cancers) and if they lack cytotoxicity, to increase the strength of the
CC immune response, particularly to a co-administered antigen
CC (tumour-associated, bacterial or viral). The method allows production of
CC mistletoe lectin, and its individual chains, in many different isoforms
CC and on a large scale, at any time of the year. Recombinant products are
CC free from toxins present in natural mistletoe extracts. This sequence
CC represents a consensus sequence of the mistletoe lectin described in the
CC specification.

XX AA: Sequence 533 AA:

Alignment_scores:
Quality: 373.00 Length: 119
Ratio: 3.885 Gaps: 1
Percent Similarity: 80.672 Percent Identity: 65.546

alignment_block:
US-09-627-165b-13 x AAY25973 ..

Align seg 1/1 to: AAY25973 from: 1 to: 533

1 GCCGAGATTCATCCCATCCTGTGGAGAGCGTGCACCAATTACAGTG 50
||||| ||||| ||||| |||||
168 AlArApheAsnProIleLeuTrParG**ArgLIn**IleAsnSerG1 184
51 GGAGCTCNITCCACCACAACATGATCATCTGTGACGTCGGAGAGCACTGGG 100
||| ||| ||||| |||||
184 Y***SerPheLeuProAsp***TyMetLeuGlutLeuGluThrSerTrpg 201
101 GTGCAGATTCACGCCACGATCGACAGCTCACAGAGTAGTCATTTTAATACG 150
||||| ||||| ||||| |||||
201 TyGlnGlnSerThrcInValGlnHisSerThrsrGlyAlaIleAsnAsn 217
151 CAATTAAGATTGCAGATTTCCGCCGCGTAACTTGTAGCGNTAGCAATGT 200
||||| ||||| ||||| |||||
218 Pro**ArgLeuAlaIle*****GlyAsnPheValAlaThrLeu**AsnVa 234
201 TCGCGAGTGATCTTCACAGCTTGCGGATCATGTTTGTCCGAATGCAGTGTG 250
||||| ||||| ||||| |||||
234 IArg***ValIleAlaSerLeuAlaIleMetLeuPheValCysGlyGluA 251
251 GGCCATTCTCTCTCTCGACACACCTTGCGCGCTGCTCTAGAGTGCCTG 300
||||| ||||| ||||| |||||
251 rProSerSerSerAspValAlaTrgTYrTPProLeuValIleArProval 267
301 GTGAGATCGCGCAACGATGTATCGATCGACTWTTCGACACCCACCGTGC 350
::: ||||| ||||| ||||| |||||
268 Ile.....AlaAspAspValThrCysSerAlaSerGluProThrValAr 282
351 CATCGTA 357
|||||
282 gIleVal 284

seq_name: /SIDSL/gcgdata/hold-geneseq/geneseq-emb./AA1999.DAT:AAY25976

seq_documentation_block:
ID: AAY25976 standard; protein: 533 AA.
XX AA: AAY25976;
XX AA: 18-OCT-1999 (first entry)
DE DE: Mistletoe lectin protein consensus sequence 3.
FM FM: Mistletoe lectin; antitumour; immunostimulant; A-chain; MLN; immunity;
RM ribosome 28S subunit; non-cytotoxic; T-cell activation; immune response;
NM lymphokine-producing macrophage; uncontrolled cell growth; treatment;
cancer; cytotoxicity; antigen; isoform.

```
XX      Viscum album.
OS      Key
XX      Location/Qualifiers
FT      Misc-difference 15 /label= Asp, Glu
FT      Misc-difference 63 /label= Gly, Glu
FT      Misc-difference 66 /label= Ile, Val
FT      Misc-difference 75 /label= Leu, Ala
FT      Misc-difference 107 /label= Asp, none
FT      Misc-difference 113 /label= Asn, Thr
FT      Misc-difference 117 /label= Pro, Thr
FT      Misc-difference 134 /label= Ser, Thr
FT      Misc-difference 141 /label= Ser, Thr
FT      Misc-difference 145 /label= Phe, Tyr
FT      Misc-difference 152 /label= Thr, Ala
FT      Misc-difference 177 /label= Ala, Tyr
FT      Misc-difference 180 /label= Tyr, Asp
FT      Misc-difference 185 /label= Ala, Glu
FT      Misc-difference 191 /label= Val, Met
FT      Misc-difference 219 /label= Ile, Phe
FT      Misc-difference 224 /label= Pro, Ser
FT      Misc-difference 225 /label= Pro, Thr
FT      Misc-difference 232 /label= Thr, Ser
FT      Misc-difference 236 /label= Asp, Ser
FT      Misc-difference 287 /label= Asn, Ser
FT      Misc-difference 290 /label= Cys, Arg
FT      Misc-difference 325 /label= Gly, Asn
FT      Misc-difference 364 /label= Gly, Asp
FT      Misc-difference 426 /label= Gly, Glu
FT      Misc-difference 435 /label= Val, Asp
FT      Misc-difference 439 /label= Glu, Lys
FT      Misc-difference 442 /label= Gly, none
FT      Misc-difference 443 /label= Arg, Lys
FT      Misc-difference 464 /label= Cys, Ser, Val
FT      Misc-difference 480 /label= Ala, Gly
FT      Misc-difference 481 /label= Gly, Ala
FT      Misc-difference 483 /label= Ser, Gly
FT      Misc-difference 484 /label= Gly, Ser
FT      Misc-difference 493 /label= Gly, Ser
```

```
FT      /label= Gly, Tyr
FT      Misc-difference 500 /label= Asn, Ser, Thr, Lys
FT      Misc-difference 501 /label= Ser, Gly
FT      Misc-difference 502 /label= Leu, Pro
FT      Misc-difference 503 /label= Ala, Met
FT      Misc-difference 504 /label= Met, Val
FT      Misc-difference 533 /label= Pro, Phe
XX
XX      DE19804210-A1.
XX
XX      12-AUG-1999.
XX
XX      03-FEB-1998; 98DE-1004210.
XX
XX      03-FEB-1998; 98DE-1004210.
XX
XX      (BIOS-) BIOSYN AR2NEIMITTEL GMBH.
XX
XX      Morris P, Stiefel T, Voelter W, Welters P;
XX      WPI, 1999-445335/38.
XX
XX      Preparation of mistletoe lectins in heterologous systems,
XX      particularly for use as anticancer agents and immunostimulants
XX
XX      Claim 40; Page 37-38; 78pp; German.
XX
XX      This invention describes a novel mistletoe lectin (I) and its fragments
XX      which have antitumour and immunostimulatory activity. The A-chain (MA)
XX      of the mistletoe lectin binds to, and inactivates, the 28S subunit of
XX      ribosomes. Non-cytotoxic forms of (I) activate T-cell and
XX      lymphokine-producing macrophages, so stimulate immunity. (I) and its
XX      fragments are used to treat uncontrolled cell growth (particularly of the
XX      cancers) and if they lack cytotoxicity, to increase the strength of the
XX      immune response, particularly to a co-administered antigen
XX      (tumour-associated, bacterial or viral). The method allows production of
XX      mistletoe lectin, and its individual chains, in many different isoforms
XX      and on a large scale, at any time of the year. Recombinant products are
XX      free from toxins present in natural mistletoe extracts. This sequence
XX      represents a consensus sequence of the mistletoe lectin described in the
XX      specification.
XX
XX      Sequence 533 AA:

alignment_scores:
    Quality: 373.00      Length: 119
    Ratio: 3.885        Gaps: 1
    Percent Similarity: 80.672      Percent Identity: 65.546

alignment_block:
US-09-627-165B-13 x MAY25976 ..

Align seg 1/1 to: MAY25976 from: 1 to: 533

1 GCCAGATTCATCCCATGCTGAGAGCTTGGCGGCAATTTACAGTGG 50
|||||
168 AlAArgPheAsnProIleuLThrPhy**ArgGln**IleAsnSerG1 184
51 GGAGTCNTCTCCACCAACATGTACATGCTGAGCTGGAGACAGTGGG 100
|||||
184 y***SerPheLeuProAsp**TyrMetLeuGluLeuGluThrSerTPG 201
101 GTGCAAAATCCACCAAGTCCACAGTCCAAAGATGGCATTTTAAATACC 150
|||||
201 TgInGlnSerThrGlnValGlnHisSerThrAspGlyValPheAsnSn 217
```


PD 12-AUG-1999.

XX

seq_name: /SIDS1/gcdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:AA25983

AC	AAV25983:
XX	
DT	18-OCT-1999 (first entry)
XX	
DE	Mistletoe lectin A1 (variant) protein fragment.
XX	
XX	Mistletoe; lectin; antitumour; immunostimulant; A-chain; M1A; immunity;
KW	ribosome 28S subunit; non-cytotoxic; T-cell activation; immune response;
KW	lymphokine-producing macrophage; uncontrolled cell growth; treatment;
KW	cancer; cytotoxicity; antigen; isoform; lectin A1.
XX	
OS	viscum album.
XX	
PN	DE19804210-A1.
XX	
PD	12-AUG-1999.
XX	
PF	03-FEB-1998; 98DE-1004210.
XX	
PR	03-FEB-1998; 98DE-1004210.
XX	
PA	(BIOS-) BIOSYN ARZNEIMITTEL GMBH.
XX	
PI	Morris P, Stiefel T, Voelter W, Welters P:
DR	WPI: 1999-445335/38.
XX	
DR	N-PSDB; AAZ09107.
XX	
PT	Preparation of mistletoe lectins in heterologous systems,
XX	particularly for use as anticancer agents and immunostimulants
XX	
PS	Disclosure: Fig 5B; 78pp; German.
XX	
XX	This invention describes a novel mistletoe lectin (I) and its fragments
CC	which have antitumour and immunostimulatory activity. The A-chain (M1A)
CC	of the mistletoe lectin binds to, and inactivates, the 28S subunit of
CC	ribosomes. Non-cytotoxic forms of (I) activate T-cell and
CC	lymphokine-producing macrophages, so stimulate immunity. (I) and its
CC	fragments are used to treat uncontrolled cell growth (particularly
CC	cancers) and if they lack cytotoxicity to increase the strength of the
CC	immune response, particularly to a co-administered antigen
CC	(tumour-associated, bacterial or viral). The method allows production of
CC	mistletoe lectin, and its individual chains, in many different isoforms
CC	and on a large scale, at any time of the year. Recombinant products are
CC	free from toxins present in natural mistletoe extracts. This sequence
CC	represents a fragment of a mistletoe lectin A1 protein variant.
XX	
XX	Sequence 254 AA:
XX	

```

alignment_scores:
  Quality: 317.00      Length: 88
  Ratio: 4.342      Gaps: 0
Percent Similarity: 82.955      Percent Identity: 71.591

alignment_block:
US-09-627-165B-13 x AAY25983      ..

Align seg 1/1 to: AAY25983 from: 1 to: 254

1 GCCAGATTCCAATCCCATNTGTGGAGGCTTGCCGCCGCAAAATTAAACAGTGG 50
167 AIAATgPheasPrCoILeUtrPaRgTYrAaGgIntYrILeasnsrcl 183
51 GGAGTCCTTCACCAACAAATGATTAATCCTCCAGCTGGAGACAGATTGGG 100
183 yAlasErPheLeuPrCoasPaRvAlTYrMeILeuILuengInuIrnsErTrpg 200
101 GTGCACAATCCACCCCAAGTCCAGCAGCTCCAGAGATGGCACTTTTAAATAC 150
200 IyGInclnsErThrgInvalGInHisSerTThAspLyAlPheasnaSn 216
151 CAATTAAGATTGCAGATTTCCCGCGGTAACCTTGTGCACNTGAGCAATGT 200

```

217 ProIleArgLeuAlaIleProProGlyAsnHeValThrLeuIrnAsnVa 233
 ||||| || | ||||| ||| ::|||
 201 mGcGACGATCTGCAGCTGGCGATCATCTTTGTCGAATGAGTC 250
 ||||| :||| ||||| |||||
 233 lArgspAlIleIleAserLeuAlaIleMetLeuPheValCysGlyAla 250
 ||||| |||||
 251 GGCCATTCCTCT 264
 ||||| |||||
 250 rgProSerSerSer 254

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:AAW64659

seq_documentation_block:

AAW64659 standard; Protein; 252 AA.

AAW64659;

DT 23-OCT-1998 (first entry)

DE Mistletoe RMLA protein.

XX lectin A-chain; mistletoe; rMLA; fusion protein; effector; cytotoxic;
 KW intracellular; processing module; protease recognition; targeting module;
 KW internalisation; treatment; disorder; cell proliferation; activation;
 KW autoimmune disease; allergy; tumour; rich; translocation.
 KW

OS Viscum album.

PN W09829540-A2.

PD 09-JUL-1998.

02-JAN-1998;
PF

PR 02-JAN-1997;

PA (BRAI-) BRAIN

PI Eck J, Schmidt

DR WPT; 1998-388

XX

PT chain - useful

XX

XX

CC This sequence

CC processing m

CC covalently b

CC protein. Suc

autoimmune d

CC at 1 ng to 5

CC cells. The p
CC fusion prote

CC non-specific

CC form that do
CC long half-11

CC it actively
CC endoplasmic

XX	Sequence
50	2

alignment_scores:

Quality: 311.00 Length: 85
Ratio: 4.380 Gaps: 0
Percent Similarity: 83.529 Percent Identity: 71.765

alignment_block:

US-09-627-165b-13 x AAM64659 ..

Align seg 1/1 to: AAM64659 from: 1 to: 252

```

1 GCCAGATTCAATCCCATCMTGTGGAGCGCTCCGCCGCAAAATTAACAGTGG 50
|||||
168 AlaArgPheAsnProIleuTrpArgAlaArgGlnTyrIleAsnSerGI 184
51 GGAGTNCWTCCACCAACATGTGCTCGAGCTGGAGAGAGAGTGGG 100
| ||| |||:|||||
184 yAlaSerPheLeuProAspValIleuLeuGlnTyrSerTrpG 201
101 GTGCAATCCACCCAAAGTCCAGCATGCCAGATGCGATTTTAATACC 150
||:|||||
201 IyGlnGlnSerThrGlnValGlnHisSerThrAspGlyValPheAsn 217
151 CAATTAAGATTGAGATTCCCGCGGTACTTGTGACAGTGCAGCAATGT 200
|||||
218 ProIleuArgLeuAlaIleProProGlyAsnPheValThrLeuThrAsn 234
201 TCGCAGATGATCTCCAGCTTGCGCATCATGTTGTCGAATGCAATGTC 250
|||||
234 IArgAspValIleAlaSerLeuAlaIleMetLeuPheValCysGlyGlu 251
251 GGCCA 255
251 rGPro 252

```

seq_name: /SIDSI/gcgcdata/hold-geneseq/geneseqp-emb1/AA1998.DAT.AAM64661

seq_documentation_block:

ID AAM64661 standard; Protein; 252 AA.

```

XX AC AAM64661;
XX DT 23-OCT-1998 (first entry)
XX DE Mistletoe rMLA variant protein.
XX KW Lectin A-chain; mistletoe; rMLA; fusion protein; effector; cytotoxic;
KW intracellular; processing module; protease recognition; targeting module;
KW internalisation; treatment; disorder; cell proliferation; activation;
KW autoimmune disease; allergy; tumour; ricin; translocation; ss.
XX OS Viscum album.
XX FT key Location/Qualifiers
XX FT Protein 1..252
XX FT /note= "partial"
XX PN WO9829540-A2.
XX PD 09-JUL-1998.
XX PF 02-JAN-1998; 98WO-EP00009.
XX PR 02-JAN-1997; 97EP-0100012.
XX PA (BRAT-) BRAIN BIOTECHNOLOGY RES & INFORMATION NE.
XX PI Eck J, Schmidt A, Zinke H;
XX DR WPI: 1998-388122/33.
XX DR N-PSDB; AAV51343.
XX

```

PT Nucleic acid encoding fusion protein containing mistletoe lectin A
PT Chain - useful for treatment of proliferative and autoimmune
PT diseases, allergies and tumours
XX
XX
XX
PS Disclosure; Fig 11a'; 115pp; German.
XX

CC This sequence encodes a variant mistletoe lectin A-chain, rMLA. This
CC sequence can be used in the construction of a fusion protein which
CC comprises an effector module that is cytotoxic intracellularly, a
CC processing module covalently bonded to the effector module and
CC containing a protease recognition sequence, and a targeting module
CC covalently bonded to the processing module, able to bind specifically to
CC the surface of a cell so as to mediate internalisation of the fusion
CC protein. Such a fusion protein can be used for treating disorders
CC involving proliferation and/or elevated activation of cells, especially
CC autoimmune disease, allergy and tumours. The proteins can be administered
CC e.g. by injection or topically but especially by intravenous injection,
CC at 1 ng to 500 mu g/kg/day, or for ex vivo use at 1 pg to 500 ng/mL.
CC Fusion proteins can develop toxic activity in a wide range of target
CC cells. The processing module prevents extracellular dissociation, and
CC fusion proteins based on mistletoe lectin A-chain are far more active
CC than those based on ricin and do have the associated problems of
CC non-specific toxicity. The protein may be expressed in a non-glycosylated
CC form that does not bind to sugar receptors in the liver, and which has a
CC long half-life in the blood, where the mistletoe lectin B-chain is used,
CC it actively assists in translocation of the ML A-chain from the
CC endoplasmic reticulum to the cytoplasm.

XX Sequence 252 AA:

alignment_scores:
Quality: 311.00 Length: 85
Ratio: 4.380 Gaps: 0
Percent Similarity: 83.529 Percent Identity: 71.765

alignment_block:

US-09-627-165b-13 x AAM64661 ..

Align seg 1/1 to: AAM64661 from: 1 to: 252

```

1 GCCAGATTCAATCCCATCMTGTGGAGCGCTCCGCCGCAAAATTAACAGTGG 50
|||||
167 AlaArgPheAsnProIleuTrpArgAlaArgGlnTyrIleAsnSerGI 183
51 GGAGTNCWTCCACCAACATGTGCTCGAGCTGGAGAGAGAGTGGG 100
| ||| |||:|||||
183 yAlaSerPheLeuProAspValIleuLeuGlnTyrSerTrpG 200
101 GTGCAATCCACCCAAAGTCCAGCATGCCAGATGCGATTTTAATACC 150
||:|||||
200 IyGlnGlnSerThrGlnValGlnHisSerThrAspGlyValPheAsn 216
151 CAATTAAGATTGAGATTCCCGCGGTAACTTGTGACAGTGCAGCAATGT 200
|||||
217 ProIleuArgLeuAlaIleProProGlyAsnPheValThrLeuThrAsn 233
201 TCGCAGATGATCTCCAGCTTGCGCATCATGTTGTCGAATGCAATGTC 250
|||||
233 IArgAspValIleAlaSerLeuAlaIleMetLeuPheValCysGlyGlu 250
251 GGCCA 255
251 rGPro 251

```


314 gllval 316

seq_name: /cgn2_6/plodata/2/1aa/6B.COMB.pep:US-08-776-059-39

seq_documentation_block:

; Sequence 39, Application US/08776059B

; Patent No. 6271368

; GENERAL INFORMATION:

; APPLICANT: LENTZEN, Hans

; APPLICANT: ECK, Jürgen

; APPLICANT: BAUR, Axel

; APPLICANT: ZINKE, Holger

; TITLE OF INVENTION: RECOMBINANT MISTLETOE LECTIN (RML)

; FILE REFERENCE: 674503-2003

; CURRENT APPLICATION NUMBER: US/08/776, 059B

; EARLIER FILING DATE: 1999-06-19

; EARLIER FILING DATE: 1996-06-25

; EARLIER FILING DATE: 1995-06-26

; NUMBER OF SEQ ID NOS: 56

; SOFTWARE: Patentl Ver. 2.0

; SEQ ID NO 39

; LENGTH: 235

; TYPE: PRT

; ORGANISM: Viscum album

US-08-776-059-39

alignment_scores:

Quality: 311.00

Ratio: 4.380

Percent Similarity: 83.529

Length: 85

Gaps: 0

Percent Identity: 71.765

US-09-627-165b-13 x US-08-776-059-39

Align seg 1/1 to: US-08-776-059-39 from: 1 to: 235

```
1 GCAGATTCAATCCATCCTGAGAGCTTCGCCGCAATTAACAGTGG 50
|||||
150 AAlarphasnProileutrpargalaarglnlyrilleasnsercl 166
51 GGAGTCNTCTCCACCAACATGACATGCTGAGAGCTGAGACGAGTGGG 100
|||||
166 yAlaserPhelauProaspvallyrmetleuglnleucluthtsertpg 183
101 GTCCAGATCCACCCAGCTGAGAGCTGAGAGATGCAATTTTAATACC 150
|||||
183 lYglnclnserThrglnvalglnhlserrThraspglyvalPhasnasn 199
151 CAATTAAGATTGCAATTCGCCGCGGTAACTTTGTGAGAGTGCATATG 200
|||||
200 ProileargleualaileProproglyasnphvalthrleutrasnva 216
201 TCCGAGATGATCTCCAGCTTGGGATCATGTTGCAATGACAGTGGT 250
|||||
216 lAraspvalillealaserleualailewetteleuphevalCysglylna 233
251 GGCA 255
233 rgPro 234
```

seq_name: /cgn2_6/plodata/2/1aa/6B.COMB.pep:US-08-776-059-31

seq_documentation_block:

; Sequence 31, Application US/08776059B

; Patent No. 6271368

; GENERAL INFORMATION:

; APPLICANT: LENTZEN, Hans

; APPLICANT: ECK, Jürgen

; APPLICANT: BAUR, Axel

; APPLICANT: ZINKE, Holger

; TITLE OF INVENTION: RECOMBINANT MISTLETOE LECTIN (RML)

; FILE REFERENCE: 674503-2003

; CURRENT APPLICATION NUMBER: US/08/776, 059B

; EARLIER FILING DATE: 1999-06-19

; EARLIER FILING DATE: 1996-06-25

; EARLIER FILING DATE: 1995-06-26

; NUMBER OF SEQ ID NOS: 56

; SOFTWARE: Patentl Ver. 2.0

; SEQ ID NO 31

; LENGTH: 253

; TYPE: PRT

; ORGANISM: Viscum album

US-08-776-059-31

alignment_scores:

Quality: 311.00

Ratio: 4.380

Percent Similarity: 83.529

Length: 85

Gaps: 0

Percent Identity: 71.765

US-09-627-165b-13 x US-08-776-059-31

Align seg 1/1 to: US-08-776-059-31 from: 1 to: 253

```
1 GCAGATTCAATCCATCCTGAGAGCTTCGCCGCAATTAACAGTGG 50
|||||
168 AAlarphasnProileutrpargalaarglnlyrilleasnsercl 184
51 GGAGTCNTCTCCACCAACATGACATGCTGAGAGCTGAGACGAGTGGG 100
|||||
184 yAlaserPhelauProaspvallyrmetleuglnleucluthtsertpg 201
101 GTCCAGATCCACCCAGCTGAGAGCTGAGAGATGCAATTTTAATACC 150
|||||
201 lYglnclnserThrglnvalglnhlserrThraspglyvalPhasnasn 217
151 CAATTAAGATTGCAATTCGCCGCGGTAACTTTGTGAGAGTGCATATG 200
|||||
218 ProileargleualaileProproglyasnphvalthrleutrasnva 234
201 TCCGAGATGATCTCCAGCTTGGGATCATGTTGCAATGACAGTGGT 250
|||||
234 lAraspvalillealaserleualailewetteleuphevalCysglylna 251
251 GGCA 255
251 rgPro 252
```

seq_name: /cgn2_6/plodata/2/1aa/5A.COMB.pep:US-08-378-761A-77

seq_documentation_block:

; Sequence 77, Application US/08378761A

; Patent No. 5635384

; GENERAL INFORMATION:

; APPLICANT: WALSH, TERENCE A

; APPLICANT: HEY, TIMOTHY D

; APPLICANT: MORGAN, ALICE BR

; TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE

; TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF

; NUMBER OF SEQUENCES: 81

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: ANDREA T. BORUCKI

; STREET: 9330 ZIONSVILLE ROAD

; CITY: INDIANAPOLIS

; STATE: IN

; COUNTRY: US

; ZIP: 46268

; COMPUTER READABLE FORM:

; MEDIUM TYPE: floppy disk

```

;
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/378,761A
; FILING DATE: 26-JAN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BORUCKI, ANDREA T
; REGISTRATION NUMBER: 33651
; REFERENCE/DOCKET NUMBER: 38272B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317) 337-4846
; INFORMATION FOR SEQ ID NO: 77:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 540 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-378-761A-77

alignment_scores:
  Quality: 177.00      Length: 119
  Ratio: 2.329        Gaps: 1
  Percent Similarity: 63.866      Percent Identity: 36.975

alignment_block:
US-09-627-165b-13 x US-08-378-761A-77  ..

Align seg 1/1 to: US-08-378-761A-77 from: 1 to: 540

1  GCCAGATTCAATCCCATGTCGTGGAGGCTTCGCCGCAATTAAACAGTGG 50
|||||
178  AlArgrhbelInyrtleGlInglumetArgrhrArgrleArgrtAr 194
51  GGAATCTCTCCACCAACATGTACATCTCGACCTGGAGACGATGGG 100
|||||
194  nArgrSerArAlArProAsPrroSerValleThrlLeuGlnAsnSerTrpG 211
101  GTCGACATTCACACCACCAAGTCACAGATTCACAGATTCGATTTTAAATACC 150
|||||
211  lYArgrleuSerThrlAlArleGlnGlnuSerAsnGlnGlyAlArpheAlaser 227
151  CAATTAAGATTCGAGATTTCCGCGTAACCTTGTGAGCGTGAAGATGT 200
|||||
228  PrroleInleuGlnArgrArgrAsnGlySerlySpheAsnVallyrAspVa 244
201  TCGCGACGTGATCTCCACGCTTGGCATCATGTGTTCGATTCGACGTGTC 250
|||||
244  lSerlleleuIleProIleleAlaIleMetVallyrArgrCysAlArProP 261
251  GCCCATTCCTCTCTCGACCAACCCCTTCGCGCGCTGCTCCTTAAGTCGTC 300
|||||
261  rOpTroSerSerGlnPheSer.....leuIleAlArgrProVal 273
301  GTGATCGCGCGCAACGATGTACATCTGACATTTTCCGACCCACCGTGG 350
|||||
274  ValProAsnPhesAlArAspValCysMetAsProGlnProIleValAr 290
351  CATGCTA 357
|||||
290  gIleVal 292

seq_name: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:US-08-485-286-77

seq_documentation_block:
; Sequence 77, Application US/08485286
; Patent No. 5646026
; Patent No. 5646026 5646119
; GENERAL INFORMATION:
; APPLICANT: WALSH, TERENCE A
```

```

;
; APPLICANT: HEY, TIMOTHY D
; APPLICANT: MORGAN, ALICE ER
; TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE
; TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESS: ANDREA T. BORUCKI
; STREET: 9330 ZIONSVILLE ROAD
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: US
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,286
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/378761
; FILING DATE: 26-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: BORUCKI, ANDREA T
; REGISTRATION NUMBER: 33651
; REFERENCE/DOCKET NUMBER: 38272B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317) 337-4846
; INFORMATION FOR SEQ ID NO: 77:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 540 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-485-286-77

alignment_scores:
  Quality: 177.00      Length: 119
  Ratio: 2.329        Gaps: 1
  Percent Similarity: 63.866      Percent Identity: 36.975

alignment_block:
US-09-627-165b-13 x US-08-485-286-77  ..

Align seg 1/1 to: US-08-485-286-77 from: 1 to: 540

1  GCCAGATTCAATCCCATGTCGTGGAGGCTTCGCCGCAATTAAACAGTGG 50
|||||
178  AlArgrhbelInyrtleGlInglumetArgrhrArgrleArgrtAr 194
51  GGAATCTCTCCACCAACATGTACATCTCGACCTGGAGACGATGGG 100
|||||
194  nArgrSerArAlArProAsPrroSerValleThrlLeuGlnAsnSerTrpG 211
101  GTCGACATTCACACCACCAAGTCACAGATTCGACGATGGCATTTTAAATACC 150
|||||
211  lYArgrleuSerThrlAlArleGlnGlnuSerAsnGlnGlyAlArpheAlaser 227
251  CAATTAAGATTCGAGATTTCCGCGTAACCTTGTGAGCGTGAAGATGT 200
|||||
228  PrroleInleuGlnArgrArgrAsnGlySerlySpheAsnVallyrAspVa 244
201  TCGCGACGTGATCTCCACGCTTGGCATCATGTGTTCGATTCGACGTGTC 250
|||||
244  lSerlleleuIleProIleleAlaIleMetVallyrArgrCysAlArProP 261
251  GCCCATTCCTCTCTCGACCAACCCCTTCGCGCGCTGCTCCTTAAGTCGTC 300
|||||
```

```
261 roProSerSerGlnPheSer.....LeuLeuIleArgProVal 273
301 GTGATTCGGCCAGCATGTCACCTGCACGCTTTTCGAAACCCACCGTGG 350
274 ValProAsnPhenAsnAlaAspValCysMetAspProGluProIleValArg 290
351 CATCGTA 357
290 GLEVAL 292

seq_name: /cgn2.6/prodata/2/1aa/5b_COMB.pep:us-08-356-786-10

seq_documentation_block:
; Sequence 10, Application US/08356786
; Patent No. 5877305
; GENERAL INFORMATION:
; APPLICANT: Huston, James S.
; APPLICANT: Oppermann, Hermann
; APPLICANT: Houston, L. L.
; APPLICANT: Ring, David B.
; TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
; TITLE OF INVENTION: Marker
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/356,786
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/831,967
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: CRP-053
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 534 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-356-786-10

alignment_scores:
Quality: 159.00 Length: 119
Ratio: 2.092 Gaps: 2
Percent Similarity: 63.866 Percent Identity: 33.613

alignment_block:
US-09-627-165B-13 x US-08-356-786-10 ..
Align seq 1/1 to: US-08-356-786-10 from: 1 to: 534

1 GCGAGATTCATCCCATCCTGTCGGAGGCTTCGCGCAATTAAGAGTGG 50
|||||
182 AlaArgPheGlnIleArgIleGluGluMetArgThrArgIleArgTyrAs 198
51 GGAATCTCTCCACCAACAGTGTACATGCTCGAGCTGGAGAGAGATTGGG 100

198 nArgArgSerAlaProAsnProSerValIleThrLeuGluAsnSerTrp 215
101 GTGCAGAAATCCAGCCAAAGTCACGACGATCCAGAGATGATGATTTAAATCC 150
|||||
215 LysArgLeuSerThrAlaIleGlnGlnIleSerAsnGlnGlnIleAlaPheAlaSer 231
151 CAAATAGATTGACAGATTTCCGCCGTAACCTTGTGACGNTGACAAATGT 200
|||||
232 ProIleGlnLeuGlnArgArgAsnGlySerIlySpheSerValTyrAspVal 248
201 TCGCAGCGTATCTCCACGCTTGGCGATCATGTGTTCGATAGCAGTGGTC 250
|
248 lSerIleLeuIleProIleIleAlaIleuMetValTyrArgCysAlaPro 265
251 GGCATCTCTCCTCTGCGACCAACCCCTTGGCGCGCTGCTCCCTTAAGTCCGC 300
|||
265 roProSerSerGlnPheSer.....LeuLeuIleArgProVal 277
301 GTGATTCGGCCAGCATGTCACCTGCACGCTTTTCGAAACCCACCGTGG 350
|||
278 ValProAsnPhenAsnAlaAspValCys.....MetAspProGluIleG 292
351 CATCGTA 357
|||
292 nLeuVal 294

seq_name: /cgn2.6/prodata/2/1aa/5a_COMB.pep:us-07-901-707-1

seq_documentation_block:
; Sequence 1, Application US/07901707
; Patent No. 5376546
; GENERAL INFORMATION:
; APPLICANT: Bernhard, Susan L.
; APPLICANT: Better, Marc D.
; APPLICANT: Cartoll, Steve F.
; APPLICANT: Lane, Julie A.
; TITLE OF INVENTION: Materials Comprising and Methods of
; TITLE OF INVENTION: Composition and Use for Ribosome-Inactivating Proteins
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; STREET: Two First National Plaza, 20 South Clark
; STREET: Street
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/901,707
; FILING DATE: 19920619
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/87,567
; FILING DATE: 04-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5376546and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27129/30910
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-5750
; TELEX: 25-3866
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 267 amino acids
; TYPE: AMINO ACID
```

```

;
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   US-07-901-707-1

alignment_scores:
    Quality: 138.00      Length: 87
    Ratio: 2.379         Gaps: 0
    Percent Similarity: 66.667      Percent Identity: 35.632

alignment_block:
US-09-627-165b-13 x US-07-901-707-1  ..

Align seg 1/1  to: US-07-901-707-1  from: 1  to: 267

1  GCCAGATTCATCCATCCATCMTGTGAGGCTTCGCCGCCAATTAACTAGTGG  50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
179  AlAaArgPheGlnTyrIleGlnGlyGluMetLarGThrArgIleArgTyrAs  195
51  GGAGTCNTCTCCACCAACATGTACATGTGCTGAGACGAGATTGGG  100
:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
195  nArgArgSerAlaProAspProSerValIleThrLeuGlnAsnSerTrpG  212
212  1yAArgLeuSerThrAlaIleGlnGluSerAsnGlnGlyAlaPheAlaSer  228
101  GTCGACATCCACCAAGTCCAGCTCCAGAGATGCGATTTTTAATACC  150
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
212  1yAArgLeuSerThrAlaIleGlnGluSerAsnGlnGlyAlaPheAlaSer  228
151  CAATTAAGATTGCGAGATTTCGCCGCTTAACCTTTGTCAGCTGAGCAATCT  200
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
229  ProIleGlnLeuGlnArgArgAsnGlySerLysPheSerValTyrAspVa  245
201  TCGGACGCTGATCTCCAGCTTGCGGATCATGTTGTCGATGCAATGCGTGC  250
|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
245  1SerIleLeuIleProIleIleAlaIleuMetValTyrArgCysAlaProp  262
251  GGCCATTCTCC  261
|||:|:|
262  rProSerSer  265

seq_name: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:US-07-988-430-1

seq_documentation_block:
; Sequence 1, Application US/07988430
; Patent No. 5416202
; GENERAL INFORMATION:
; APPLICANT: Bernhard, Susan L.
; APPLICANT: Better, Marc D.
; APPLICANT: Carroll, Stephen F.
; APPLICANT: Lane, Julie A.
; APPLICANT: Lei, Shau-Ping
; TITLE OF INVENTION: Materials Comprising and Methods of
; TITLE OF INVENTION: Preparation and Use for Ribosome-Inactivating Proteins
; NUMBER OF SEQUENCES: 101
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; STREET: Two First National Plaza, 20 South Clark
; STREET: Street
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/988,430
; FILING DATE: 19921209
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/901,707
```

```

;
;   FILING DATE: 19-JUN-1992
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/787,567
;   FILING DATE: 04-NOV-1991
;   ATTORNEY/AGENT INFORMATION:
;   NAME: No. 5416202and, Greta E.
;   REGISTRATION NUMBER: 35302
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (312) 346-5750
;   TELEFAX: (312) 984-9740
;   TELEX: 25-3856
;   INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 267 amino acids
;   TYPE: AMINO ACID
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   US-07-988-430-1

alignment_scores:
    Quality: 138.00      Length: 87
    Ratio: 2.379         Gaps: 0
    Percent Similarity: 66.667      Percent Identity: 35.632

alignment_block:
US-09-627-165b-13 x US-07-988-430-1  ..

Align seg 1/1  to: US-07-988-430-1  from: 1  to: 267

1  GCCAGATTCATCCATCCATCMTGTGAGGCTTCGCCGCCAATTAACTAGTGG  50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
179  AlAaArgPheGlnTyrIleGlnGlyGluMetLarGThrArgIleArgTyrAs  195
51  GGAGTCNTCTCCACCAACATGTACATGCTGAGCTGAGACGAGATTGGG  100
:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
195  nArgArgSerAlaProAspProSerValIleThrLeuGlnAsnSerTrpG  212
101  GTCGACATCCACCAAGTCCAGCTCCAGAGATGCGATTTTTAATACC  150
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
212  1yAArgLeuSerThrAlaIleGlnGluSerAsnGlnGlyAlaPheAlaSer  228
151  CAATTAAGATTGCGAGATTTCGCCGCTTAACCTTTGTCAGCTGAGCAATGT  200
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
229  ProIleGlnLeuGlnArgArgAsnGlySerLysPheSerValTyrAspVa  245
201  TCGGACGCTGATCTCCAGCTTGCGGATCATGTTGTCGATGCAATGCGTGC  250
|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
245  1SerIleLeuIleProIleIleAlaIleuMetValTyrArgCysAlaProp  262
251  GGCCATTCTCC  261
|||:|:|
262  rProSerSer  265

seq_name: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:US-08-218-303-16

seq_documentation_block:
; Sequence 16, Application US/08218303
; Patent No. 5547867
; GENERAL INFORMATION:
; APPLICANT: Kara, Bhupendra V.
; APPLICANT: Hockney, Robert C.
; APPLICANT: Fitton, John B.
; TITLE OF INVENTION: FERMENTATION PROCESS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cushman, Darby & Cushman
; STREET: 1615 L Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20036-5601
```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/218,303
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/841,533
FILING DATE: 26-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Kokulis, Paul N.
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: FPK/3893/94908/MJW
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3000
TELEFAX: 202-822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 267 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-218-303-16

alignment_scores:
Quality: 138.00      Length: 87
Ratio: 2.379         Gaps: 0
Percent Similarity: 66.667   Percent Identity: 35.632

alignment_block:
US-09-627-165b-13 x US-08-218-303-16  ..

Align seg 1/1  to: US-08-218-303-16  from: 1  to: 267

1  GCCGAGATTCAATCCCATGTCGAGCGCTGCCGCCAATTATACAGTGG  50
|||||
179  AlArghPhgIntyrlleGluglYlUmeLArghrArghleArghyrAs  195
::: ||| |||::: |||::: |||::: |||::: |||::: |||::: |||
195  nArghSerAlArProAspProSerValIlethrlengluAnsSerTrpG  212
101  GTCGACAAATCCACCAAGTCACGAGTCGAGAGATGGATTTTAATACC  150
|||||
212  lYArghLeuSerThrAlaIleGlnGlnSerAnGlnGlnAlaPheAlaSer  228
151  CAAATTAAGATTGCAGATTTCGCCGCGTAACCTTGTGACGNTGAGCAATG  200
|||||
229  ProIleGlnLeuGlnArghArghAsnGlySerIspheSerValTyrrAspA  245
201  TCGGAGAGTATCTCCAGCTTGCGCATGATGTGTGCAATGAGAGTGATC  250
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
245  lSerIleLeuIleProIleIleAlaLeuMetValTyrrGcysAlaProP  262
251  GGCCATTCTCC  261
|||
262  rProSerSer  265

seq_name: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:US-08-425-336-1

seq_documentation_block:
: Sequence 1, Application US/08425336
: Patent No. 5621083
: GENERAL INFORMATION:
: APPLICANT: Better, Marc D.
: APPLICANT: Carroll, Stephen F.
: APPLICANT: Studilka, Gary M.
: TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
```

```

TITLE OF INVENTION: Proteins
NUMBER OF SEQUENCES: 140
CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6100 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/425,336
FILING DATE: 18-APR-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/064,691
FILING DATE: 12-MAY-1993
APPLICATION NUMBER: US 07/901,707
FILING DATE: 19-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/787,567
FILING DATE: 04-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Thomas C.
REGISTRATION NUMBER: P-36,989
REFERENCE/DOCKET NUMBER: 31394
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 267 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-425-336-1

alignment_scores:
Quality: 138.00      Length: 87
Ratio: 2.379         Gaps: 0
Percent Similarity: 66.667   Percent Identity: 35.632

alignment_block:
US-09-627-165b-13 x US-08-425-336-1  ..

Align seg 1/1  to: US-08-425-336-1  from: 1  to: 267

1  GCCGAGATTCAATCCCATGTCGAGCGCTGCCGCCAATTATACAGTGG  50
|||||
179  AlArghPhgIntyrlleGluglYlUmeLArghrArghleArghyrAs  195
::: ||| |||::: |||::: |||::: |||::: |||::: |||::: |||
195  nArghSerAlArProAspProSerValIlethrlengluAnsSerTrpG  212
101  GTCGACAAATCCACCAAGTCACGAGTCGAGAGATGGATTTTAATACC  150
|||||
212  lYArghLeuSerThrAlaIleGlnGlnSerAnGlnGlnAlaPheAlaSer  228
151  CAAATTAAGATTGCAGATTTCGCCGCGTAACCTTGTGACGNTGAGCAATG  200
|||||
229  ProIleGlnLeuGlnArghArghAsnGlySerIspheSerValTyrrAspA  245
201  TCGGAGAGTATCTCCAGCTTGCGCATGATGTGTGCAATGAGAGTGATC  250
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
245  lSerIleLeuIleProIleIleAlaLeuMetValTyrrGcysAlaProP  262
```

251 GGCATCTCC 261
||| |||
262 roProSer 265

seq_name: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:US-08-488-113B-1

seq_documentation_block:

Sequence 1, Application US/08488113B

Patent No. 5744580

GENERAL INFORMATION:

APPLICANT: Better, Marc D.

APPLICANT: Carroll, Stephen F.

APPLICANT: Studnika, Gary M.

TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating

TITLE OF INVENTION: Proteins

NUMBER OF SEQUENCES: 169

CORRESPONDENCE ADDRESS:

ADDRESSEE: McAndrews, Held & Malloy, Ltd.

STREET: 500 West Madison Street, 34th floor

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60661

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/488,113B

FILING DATE: 07-JUN-1995

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/425,336

FILING DATE: 18-APR-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/064,691

FILING DATE: 12-MAY-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/988,430

FILING DATE: 09-DEC-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/901,707

FILING DATE: 19-JUN-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/787,567

FILING DATE: 04-NOV-1991

ATTORNEY/AGENT INFORMATION:

NAME: McNicholas, Janet M.

REGISTRATION NUMBER: 32,918

REFERENCE/DOCKET NUMBER: 110220S07/200-70.P3.C2A

TELEPHONE: 312/707-8889

TELEFAX: 312/707-9155

TELEX: 650 388-1248

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 267 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-488-113B-1

alignment_scores:

Quality: 138.00 Length: 87

Ratio: 2.379 Gaps: 0

Percent Similarity: 66.667 Percent Identity: 35.632

alignment_block:

US-09-627-165b-13 x US-08-488-113B-1

Align seg 1/1 to: US-08-488-113B-1 from: 1 to: 267

seq_name: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:US-08-477-484B-1

seq_documentation_block:

Sequence 1, Application US/08477484B

Patent No. 5756699

GENERAL INFORMATION:

APPLICANT: Better, Marc D.

APPLICANT: Carroll, Stephen F.

APPLICANT: Studnika, Gary M.

TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating

TITLE OF INVENTION: Proteins

NUMBER OF SEQUENCES: 169

CORRESPONDENCE ADDRESS:

ADDRESSEE: McAndrews, Held & Malloy, Ltd.

STREET: 500 West Madison Street, 34th floor

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60661

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/477,484B

FILING DATE: 07-JUN-1995

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/425,336

FILING DATE: 18-APR-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/064,691

FILING DATE: 12-MAY-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/988,430

FILING DATE: 09-DEC-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/901,707

FILING DATE: 19-JUN-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/787,567

FILING DATE: 04-NOV-1991

ATTORNEY/AGENT INFORMATION:

NAME: McNicholas, Janet M.

REGISTRATION NUMBER: 32,918

REFERENCE/DOCKET NUMBER: 110220S07/200-70.P3.C2A

TELEPHONE: 312/707-8889

TELEFAX: 312/707-9155

TELEX: 650 388-1248

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 267 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-488-113B-1


```

APPLICANT: Barth, Peter Thomas
TITLE OF INVENTION: VECTOR
NUMBER OF SEQUENCES: 61
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN DABRY CUSHMAN
ADDRESSEE: INTELLECTUAL PROPERTY GROUP OF
ADDRESSEE: PILLSBURY MADISON & SUTRO, L.L.P.
STREET: 1100 New York Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
COMPUTER: IBM PC/XT/AT Compatibles
OPERATING SYSTEM: MS-DOS
SOFTWARE: Microsoft Word or ASCII editors
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/338, 793D
FILING DATE: 08-No. 5840521-94
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/842, 081
FILING DATE: 26-Feb-92
CLASSIFICATION: 435
APPLICATION NUMBER: 9104017.0
FILING DATE: 26-Feb-91
APPLICATION NUMBER: 9109188.4
FILING DATE: 29-Apr-91
ATTORNEY/AGENT INFORMATION:
NAME: Kokulis, Paul N.
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: DDB/9901/215431/TCW
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3000
TELEFAX: 202-822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 61:
SEQUENCE CHARACTERISTICS:
LENGTH: 267 amino acids
TYPE: amino acid
STRANDEDNESS: Single
TOPOLOGY: Linear
US-08-338-793D-61

alignment_scores:
Quality: 138.00 Length: 87
Ratio: 2.379 Gaps: 0
Percent Similarity: 66.667 Percent Identity: 35.632

alignment_block:
US-09-627-165b-13 x US-08-338-793D-61 ..

Align seg 1/1 to: US-08-338-793D-61 from: 1 to: 267

1 GCGAGATTCAATCCATCTGTCGAGCGCTCGCGCAATTAACAGTGG 50
|||||
179 Alaaagphgintyrlleglmglyglumelargthrarqileatgyras 195
51 GGAGTCTCTCCACCAACATATACATGCTGAGCTGGAGACGAGTGGG 100
|||||
195 nargargseralaproasploserValilehrleuInsnsertpg 212
101 GTGGAGATTCACCCCAAGTCACAGATCCAGATGCGATTGTAATACC 150
|||||
212 lYargleuserthralaileglnluserAanglnclYAlaphealaser 228
151 CAATAAGATTGCAGATTTCGCCCGGTAACTTTGGACGNTGAGCAATCT 200
|||||
229 ProileglnleuInslnargargnsnglyserLyspheserValtyraspa 245
201 TCGGAGCGTATCTCCAGCTTGCGATCATGTTGTCGAATCAGTGTCTC 250

```

```

seq_name: /cgn2_6/ptodata/2/1aa/5b_COMB.pep:us-08-839-765-1
seq_documentation_block:
; Sequence 1, Application US/08839765
; Patent No. 614631
; GENERAL INFORMATION:
; APPLICANT: Better, Marc D.
; APPLICANT: Carroll, Stephen F.
; APPLICANT: Stuhlika, Gary M.
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
; NUMBER OF SEQUENCES: 169
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.
; STREET: 500 West Madison Street, 34th floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60661
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/839,765
; FILING DATE: 15-APR-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/425,336
; FILING DATE: 18-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/064,691
; FILING DATE: 12-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/988,430
; FILING DATE: 09-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/901,707
; FILING DATE: 19-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/787,567
; FILING DATE: 04-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McNicholas, Janet M.
; REGISTRATION NUMBER: 32,913
; REFERENCE/DOCKET NUMBER: 110220509/200-70.P3.C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/707-8889
; TELEFAX: 312/707-9155
; TELEX: 650 388-1248
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 267 amino acids
; TYPE: amino acid
; MOLECULE TYPE: linear
; US-08-839-765-1

alignment_scores:
Quality: 138.00 Length: 87
Ratio: 2.379 Gaps: 0
Percent Similarity: 66.667 Percent Identity: 35.632

alignment_block:

```

US-09-627-165B-13 x US-08-839-765-1 ..

Align seg 1/1 to: US-08-839-765-1 from: 1 to: 267

[illegible]

OM of: US-09-627-165B-13 to: PIR_71:* out_format : pts

About: Results were produced by the GenCore software, version 4.5.
Copyright (c) 1993-2000 CompuGen Ltd.

```

-MODEL=firme.np.model -DEV=xlp
-O/cgcn2.1/USPO.spool/USO9671165/runal_30072002_151139_3235/app_query.fasta_1.1012
-DB=pic.71 -PEMT=fastin -SUFIX=trp -GAPOF=12.000 -GAPEXT=4.000
-MIMMATH=0.100 -LOOPLY=0.000 -LOOPEXT=0.000 -CGAPOF=4.500
-OGAPEXT=0.050 -XGAPOF=10.000 -FGAPOF=6.000
-FGAPEXT=7.000 -YGAPOF=10.000 -XGAPEXT=0.500 -DELOP=6.000
-DELETEXT=7.000 -START=1 -MATRIX=blomsu62 -RRANS=humano40.cdl
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pls -NORM=ext -HEAPSIZE=500
-MINLEN=0 -MAXLEN=2000000000 -USER=USO9671165_tcgcn1_1.172
-NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEC=130
-NO_XLPLY -WAIT -THREADS=1

```

```
Query: US-03-02, 1000 10
Query length: 357
Database: PIR-71:*
Database sequences: 283138
Database length: 96089334
Search time (sec): 72.560000
```

Sequence	Sturd	Orig	ZScore	EScore	Len	Document
pr1:2:PCD018	+	317.00	319.58	8.0e-27	254	1
pr1:2:RUC018	+	178.00	63.6-11	2.6e-11	576	1
pr1:2:RLCSAG	+	177.00	332.81	3.4e-11	554	1
pr1:2:WZLSA	+	148.00	275.48	5.6e-08	528	1
pr1:2:S24331	+	138.00	255.50	7.3e-07	562	1
pr1:2:S32430	+	138.00	254.94	7.4e-07	528	1
pr1:2:S16222	+	135.00	249.53	1.6e-06	527	1
pr1:2:S62627	+	135.00	221.86	0.055	570	1
pr1:2:S839761	+	121.50	192.99	0.0110	106	1
pr1:2:CG39761	+	99.50	153.28	0.0584	251	1
pr1:2:RLPUG6	+	93.50	169.12	0.0875	286	1
pr1:2:R113535	+	92.00	158.79	0.3171	300	1
pr1:2:S28421	+	87.00	152.80	0.3171	294	1
pr1:2:CG87585	-	84.00	152.80	0.6817	212	1
pr1:2:JEO401	-	79.00	145.83	2.34	261	1
pr1:2:JEO5032	+	79.00	142.47	3.09	247	1
pr1:2:CG5606	+	76.50	138.06	4.64	289	1
pr1:2:U00393	+	76.00	138.47	5.16	247	1
pr1:2:S48002	+	76.00	128.47	6.09	754	1
pr1:2:S52444	+	74.00	133.45	8.76	277	1
pr1:2:S17757	+	74.00	132.36	8.92	730	1
pr1:2:TA33963	+	74.00	124.77	10.12	822	1
pr1:2:TC02824	+	73.50	122.71	11.70	822	1
pr1:2:113177	+	72.50	119.62	13.77	1038	1
pr1:2:TW1802	-	72.50	136.89	11.55	135	1
pr1:2:HB4355	-	72.50	130.75	12.79	268	1
pr1:2:RLMTZ	+	72.50	130.07	12.94	289	1
pr1:2:TW34649	+	71.50	133.83	15.19	152	1
pr1:2:AB87247	+	71.00	117.99	22.10	797	1
pr1:2:TA44066	+	70.00	116.78	28.18	730	1
pr1:2:TA61221	+	69.50	112.49	33.84	1054	1
pr1:2:U00146	+	69.00	117.60	34.74	533	1
pr1:2:BG97071	+	69.00	114.27	36.72	773	1
pr1:2:S30135	+	68.50	122.76	35.65	268	1
pr1:2:J07053	+	68.50	121.28	36.53	316	1
pr1:2:S231256	-	68.50	118.12	38.51	450	1
pr1:2:R23821	+	68.00	122.35	40.12	251	1
pr1:2:ASLDFP	+	68.00	121.93	40.40	263	1
pr1:2:S59430	+	68.00	116.79	44.01	467	1
pr1:2:S26191	-	68.00	116.67	44.10	473	1

seq_name: plr2:PD0018

C:Species: Viscum album
C:Date: 03-Jul-1998 #sequence_revision 10-Jul-1998 #text_change 26-Aug-1999
C:Accession: PD0018
R:Eschenbuerger, S.; Krauspenhaar, R.; Mikhailov, A.; Stoeva, S.; Betzel, C.; Voelter, M.
R.Biochem. Biophys. Res. Commun. 247, 367-372, 1998
A:Title: Primary structure and molecular modeling of mistletoe lectin I from Viscum album
A:Reference number: PD0018; M0ID:98308123
A:Accession: PD0018
A:Molecule type: protein
A:Residues: 1-254 <ESCS>
C:Superfamily: ricin; rRNA N-glycosidase homology
E:1-246/Domain: rRNA N-glycosidase homology <RNG>

alignment_block:
US-09-627-165B-13 x PD0018

```

1  GCCAGATTCAATCCCATCNCNTGGAGGCTTGGCCGCCAAATTACAGTGG  50
   |||||
167  A|Aa|gPheaa|P|P|Leu|Thr|P|Arg|G|I|n|T|y|L|e|a|n|S|e|r|G|  183
   |||||
51  GGAGTCNTCTCCACC|CA|A|A|C|A|T|G|A|C|T|G|C|T|G|A|G|C|T|G|G|A|N|C|A|G|T|T|G|G  100
   |||||
183  y|A|S|e|r|P|H|e|u|P|r|o|S|a|P|a|I|y|T|y|M|e|T|L|e|u|G|I|n|G|I|n|H|S|e|r|T|P|S  200
   |||||
101  G|T|G|C|A|A|A|T|C|C|A|C|C|A|G|T|C|C|A|G|C|A|G|T|C|C|A|G|A|T|G|G|C|A|T|T|T|T|A|A|T|A|C  150
   |||||
200  I|y|G|I|n|S|e|r|T|H|r|G|I|n|A|I|G|I|H|S|e|r|T|H|S|P|G|I|y|V|A|I|P|H|e|a|n|A|N  216
   |||||
151  C|A|A|T|A|A|G|A|T|T|C|G|A|T|T|C|G|C|G|G|T|A|C|T|T|G|T|G|A|C|G|N|T|G|A|C|A|T|G|T  200
   |||||
217  P|H|I|A|I|A|I|G|I|e|u|A|I|H|P|P|H|I|D|I|y|S|N|P|H|e|V|A|I|H|I|R|e|u|T|H|A|N|V|A  233
   |||||
201  T|T|C|G|G|A|G|T|A|T|C|C|A|G|C|T|T|G|G|C|A|T|C|A|T|G|T|T|G|T|C|A|A|T|G|C|A|G|T|G|T|C  250
   |||||
233  I|A|N|G|A|P|V|A|I|I|A|I|S|e|r|I|e|u|A|I|A|I|e|W|e|T|H|e|u|P|H|e|V|A|I|C|y|S|G|I|y|I|A  250
   |||||
251  G|G|C|A|T|T|T|C|T|C|T|C|T  264
   |||||
250  I|g|P|r|o|S|e|r|S|e|r|  254

```

```

seq_name: p1r1:R1CSD
seq_documentation_block:
  ricin D precursor - castor bean
N:Contains: rRNA N-glycosidase (EC 3.2.2.22)
C:Species: Ricinus communis (castor bean)
C:date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jul-1999
C:accession: A24041; S20513; A24614; A03372; A24010; A03374; S10903
R:Halling, K.C.; Halling, A.C.; Murray, E.E.; Ladin, B.F.; Houston, L.L.; Weaver, R.F
Nucleic Acids Res. 13, 8019-8033, 1985
A:title: Genomic cloning and characterization of a ricin gene from Ricinus communis
A:reference_number: A24041; M01D:86067214
A:accession: A24041
A:molecule_type: DNA
A:residues: 1-576 <HML>

```


F:35-281/Domain: rRNA N-glycosidase homology <RNG>
 F:303-564/Product: agglutinin chain B #status experimental <BCH>
 F:319-361/Region: 40-residue repeats
 F:324-402/443-450-485,489-528 #status experimental
 F:324-259/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:104,147,231,233/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted
 F:200,203/Active site: Glu, Arg #status predicted
 F:382-306,322-341,365-382,453-466,492-509/Disulfide bonds: #status predicted
 F:324,337,348/Binding site: N-acetylglucosamine (Asp, Glu, Asn) #status predicted
 F:397,437/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:536,557/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted

alignment_scores:

Quality: 177.00 Length: 119
 Ratio: 2.329 Gaps: 1
 Percent Similarity: 63.866 Percent Identity: 36.975

alignment_block:

US-09-627-165B-13 x RLCSAG ..

Align seg 1/1 to: RLCSAG from: 1 to: 564

```

1 GCCAGATTCATCCCATCNCNTGAGCGCTTGGCCGCAATTACAGTGG 50
|||||.....|||.....|||.....|||.....|||.....|||
202 AAlaArgPheGlnTyrIleGluGlyGluMetArgThrArgIleArgTyrAs 218
51 GGAGCTCNCCTCCACCAACATGTACATGCTGAGCTGGAGAGAGTGGG 100
::: ||| |||::: |||::: |||::: |||::: |||::: |||
218 nArgArgSerAlaProAspProSerValIleThrLeuGlnAsnSerTrpG 235
101 GTGCAATTCACCCCAATGTCAGAGTCCAGAGATGCGATTTCATTAATAC 150
|||||.....|||.....|||.....|||.....|||.....|||
235 IyArgLeuSerThrAlaIleGlnGluSerAsnGlnGlyAlaPheAlaSer 251
151 CAATTAAGATGCGAGATTTCGCGCGTACCTTGGACGNCAGACAGTGT 200
|||||.....|||.....|||.....|||.....|||.....|||
252 ProlIeGlnLeuGlnArgAsnGlySerIySphaAsnValIyTyrAspVa 268
301 GTGGATGGGGCCACGATGTCACGTCGACTTTTCCGACCAACCGCTGCG 350
|||||.....|||.....|||.....|||.....|||.....|||
298 ValProAsnPheAsnAlaAspValTyrMetAspProGluProIleValAla 314
351 CATCGTA 357
|||||
314 gIleVal 316
```

seq_name: p1r1:72LSA

seq_documentation_block:

abrin-a precursor - Indian licorice (fragment)
 N:Contains: rRNA N-glycosidase (EC 3.2.2.22)
 C:Species: Abrus precatorius (Indian licorice)
 C:Date: 31-Dec-1993 #sequence, revision 01-Aug-1997 #text, change 16-Jul-1999
 C:Accession: S32429; J70202; A39761; JCI398; S14472; S24133; S74110; S74111
 R:Hung, C.H.; Lee, M.C.; Lee, T.C.; Lin, J.Y.
 J. Mol. Biol. 229, 263-267, 1993
 A:Title: Primary structure of three distinct isoabrin determined by cDNA sequencing. CD
 A:Reference number: S32429; MUID:93132798
 A:Accession: S32429
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 'E', 2-528 <HUN>
 A:Cross-references: GB:M98344; NID:q166294; PIDN:AAA3624.1; PID:q166295
 A:Note: the coding region for the sequence shown is preceded by an ATG codon
 A:Note: residues 1-8 were derived from the synthesized primer

R:Funatsu, G.; Taguchi, Y.; Kamenosono, M.; Yanaka, M.

Agric. Biol. Chem. 52, 1095-1097, 1988

A:Title: The complete amino acid sequence of the A-chain of abrin-a, a toxic protein

A:Reference number: J70202

A:Accession: J70202

A:Molecule type: protein

A:Residues: 1-201,203-251 <FUN>

A:Note: The amino-terminal residue forms pyrrolidone carboxylic acid; therefore, we h

R:Evensen, G.; Mathiesen, A.; Sundan, A.

J. Biol. Chem. 266, 6848-6852, 1991

A:Title: Direct molecular cloning and expression of two distinct abrin A-chains.

A:Reference number: A39761; MUID:91201329

A:Accession: A39761

A:Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 'E', 2-251 <EVE>

A:Cross-references: GB:X54872

A:Note: residues 1-8 were derived from the synthesized primer

R:Kimura, M.; Sumizawa, Y.; Funatsu, G.

Biosci. Biotechnol. Biochem. 57, 166-169, 1993

A:Title: The complete amino acid sequences of the B-chains of abrin-a and abrin-b, to

A:Reference number: JCI398; MUID:93169023

A:Contents: seeds

A:Accession: JCI398

A:Molecule type: protein

A:Residues: 261-347, 'T', 349-351, 'A', 353-357, 'L', 359-528 <KIM>

A:Experimental source: seed

R:Evensen, G.; Mathiesen, A.; Sundan, A.

submitted to the EMBL Data Library, October 1990

A:Description: Direct molecular cloning of two distinct abrin A-chains.

A:Reference number: S14471

A:Accession: S14472

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 'M', 2-251 <EV2>

A:Cross-references: EMBL:X54873; NID:q16090; PIDN:CAA3655.1; PID:q16091

R:Chen, Y.L.; Chow, L.P.; Tsugita, A.; Lin, J.Y.

FEBS Lett. 309, 115-118, 1992

A:Title: The complete primary structure of abrin-a B chain.

A:Reference number: S24133; MUID:92371656

A:Accession: S24133

A:Molecule type: protein

A:Residues: 262-297, 'Y', 299-426, 'L', 428-466, 'P', 468-482, 'L', 484-528 <CHE>

R:Lin, S.H.; Chow, L.P.; Chen, Y.L.; Liaw, Y.C.; Chen, J.K.; Lin, J.Y.

Eur. J. Biochem. 240, 564-569, 1996

A:Title: Probing the domain structure of abrin-a by tryptic digestion.

A:Reference number: S74110; MUID:97008945

A:Accession: S74110

A:Molecule type: protein

A:Residues: 89-108,154-172 <LIN>

A:Experimental source: seed

A:Accession: S74111

A:Molecule type: protein

A:Residues: 262-276, 'X', 278-280, 329-348, 369-388, 399-418 <LIN>

A:Experimental source: seed

C:Comment: Abirin-a is more toxic than ricin. The toxin consists of an A chain, which

taining receptors on the cell surface. The A and B chains are linked by a single disu

C:Superfamily: ricin; rRNA N-glycosidase homology

C:Keywords: duplication; glycoprotein; glycosidase; hydrolase; lectin; pyroglyutamic a

F:1-251/Product: abrin-a chain A #status experimental <ACH>

F:1-246/Domain: rRNA N-glycosidase homology <RNG>

F:261-528/Product: abrin-a chain B #status experimental <BCH>

F:283-325,326-366,369-407,414-449,453-492,495-528/Region: 40-residue repeats

F:2/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F:14,113,195,196/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted

F:164,167/Active site: Glu, Arg #status predicted

F:247-269,286-305,329-346,417-430,456-473/Disulfide bonds: #status predicted

F:288,312/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted

F:361,401/Binding site: carbohydrate (Asn) (covalent) #status experimental

F:500,521/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted

alignment_scores:

```

seq_documentation_block:
  abrin-d precursor - Indian licorice (fragment)
  N:Contains: rRNA N-glycosidase (EC 3.2.2.22)
  C:Species: Abrus precatorius (Indian licorice)
  C:Date: 30-Sep-1993 #sequence_revision 01-Aug-1997 #text_change 01-Aug-1997
  C:Accession: S32431; S34408
  R:Hung, C.H.; Lee, M.C.; Lee, T.C.; Lin, J.Y.
  J. Mol. Biol. 229, 263-267, 1993
  A:Title: Primary structure of three distinct isoabryns determined by cDNA sequencing. CC
  A:Reference number: S32429; MUID:93132798
  A:Accession: S32431
  A:Molecule type: mRNA
  A:Residues: 1-528 <HUN>
  A:Cross-references: GB:M98346
  R:Hung, C.; Lee, M.; Lee, T.; Lin, J.
  submitted to the EMBL Data Library, March 1993
  A:Reference number: S34408
  A:Accession: S34408
  A:Molecule type: mRNA
  A:Residues: 1-169, 'C', 171-320, 'L', 322-528 <HU2>
  A:Cross-references: GB:M98346
  C:Comment: Abrin consists of an A chain, which inhibits protein synthesis by inactivating
  C:Superfamily: richn, rRNA N-glycosidase homology
  C:Keywords: disulfide bond; duplication; glycoprotein; glycosidase; hydrolase; lectin; F
  F:1-251/Product: abrin-d chain A #status predicted <ACH>
  F:1-246/Domain: rRNA N-glycosidase homology <RMG>
  F:561-528/Product: abrin-d chain B #status predicted <BCH>
  F:283-325-326-366, 369-407, 414-449, 453-492, 495-528/Region: 40-residue repeats
  F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status predicted

```

F;35-285/Product: abrin-c chain A #status predicted <ACH>

F;41-280/Domain: rRNA N-glycosidase homology <RNG>

F:295-562/Product: abrin-c chain B #status predicted <BCB>
 F:317-359, 360-400, 403-441, 448-483, 487-526, 529-562/Region: 40-residue repeats
 F:33/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
 F:108,147,229,230/Binding site: substrate (Tyr, Tyr, Gln, Asn) #status predicted
 F:198,201/Active site: Gln, Arg #status predicted
 F:234,287,395,435,436/Binding site: carbonyl (Asn) (covalent) #status predicted
 F:281-303,320-339,363-380,451-464,490-507/Disulfide bonds: #status predicted
 F:322,346/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted
 F:534,555/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted

alignment_scores: Length: 119
 Quality: 138.00 Gaps: 4
 Ratio: 1.792 Percent Identity: 36.134
 Percent Similarity: 64.706

alignment_block:
 US-09-627-165b-13 x S16022 ..

Align seg 1/1 to: S16022 from: 1 to: 562

```

1  GCCAGATTCATCCATCCTGAGAGCTTCGCCGCAATTACAGTGG 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
200  AAlargyrrArgrrrYrrrllleSerAsnArgValGlyValSerlleArgrrrG 216
51  GGAGTCCTCCACCAACATGATGCTCGAGCTGAGAGAGTGG 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||
216  YThAlrPheGlnProAspProAlaMetleuSerleuGlnAsnAsnTPA 233
101  GTGAGCATATCCACCAAGTCCAGCATCCAGATGGCATTTT..AAT 147
|||||:|||||:|||||:|||||:|||||:|||||:|||||
223  sPAnleuSerGlyGlyValGlnGlnSerValGlnAspThrProAsn 249
148  ACCCAATAAAGATTCAGATTCGCCGCTGACTTGTGACGNTGACAA 197
|||||:|||||:|||||:|||||:|||||:|||||:|||||
250  AsnVallleuSerSerlleAsnValGlnProValValAspSerle 266
198  TGTTCGAGCATGATCTCCAGCTTCGAGCATCTGTCGATGAGTGG 247
|||||:|||||:|||||:|||||:|||||:|||||:|||||
266  uSerHlSPrrHrValAlaValAlaValleuAlaMetleuPheValCysAsn 282
248  GTGCGCATTCCTCTCTCTCGACCACTTCGCGCTGCTCTCAAGTCC 297
|||||:|||||:|||||:|||||:|||||:|||||:|||||
283  ....ProProAsnAla.....AsnGlnSerProleuLeuIleArgSer 295
298  GTGCTGATGCGCGCAACGATGCTGACCTGCTTTTCCGAAACCCAGCT 347
|||||:|||||:|||||:|||||:|||||:|||||:|||||
296  lleValGlnGlnSerlys...lleCysSerSerArgrrrGlnProHrVa 311
348  GCGCATC 354
|||||:|||||:|||||:|||||:|||||:|||||:|||||
311  lArglle 313

```

seq_name: p1r2:S32430

seq_documentation_block:
 abrin-b precursor - Indian licorice (fragment)
 N:Contains: rRNA N-glycosidase (EC 3.2.2.22)
 C:Species: Adrus precatorius (Indian licorice)
 C:Date: 30-Sep-1993 #sequence_revision 01-Aug-1997 #text_change 20-Aug-1999
 C:Accession: S32430; JCI399
 R:Hung, C.H.; Lee, M.C.; Lee, T.C.; Lin, J.Y.
 J.Mol. Biol. 229, 263-267, 1993
 A:Title: Primary structure of three distinct isoabirins determined by cDNA sequencing.
 A:Reference numbers: S32429; M01D:93132798
 A:Accession: S32430
 A:Molecule type: rRNA
 A:Residues: 1-527 <RNN>
 A:Cross-references: GB:M98345; NID:g166296; P1DN:AA32625.1; PID:g166297
 R:Kimura, M.; Sumizawa, T.; Funatsu, G.
 Biosci. Biotechnol. Biochem. 57, 166-169, 1993
 A:Title: The complete amino acid sequences of the B-chains of Abirin-a and Abirin-b, toxic
 A:Reference numbers: JCI398; M01D:93169023

A:Accession: JCI399
 A:Molecule type: protein
 A:Residues: 260-281, 'D', 283-290, 'N', 292-349, 'PQ', 352-377, 'N', 379-425, 'W', 427, 'D', 429-
 A:Experimental source: seed
 C:Superfamily: rlcIn: rRNA N-glycosidase homology
 C:Keywords: disulfide bond; duplication; glycoprotein; glycosidase; hydrolase; lectin
 F:1-230/Product: abrin-b chain A #status predicted <BCH>
 F:7-245/Domains: rRNA N-glycosidase homology <RNG>
 F:260-527/Product: abrin-b chain B #status experimental <BCB>
 F:282-324,325-365,368-400,413-448,452-491,494-527/Region: 40-residue repeats
 F:1/Modified site: pyrrolidone carboxylic acid (Gln) (covalent) #status predicted
 F:110,136,400/Binding site: substrate (Tyr, Tyr, Gln, Asn) #status predicted
 F:163,166/Active site: Gln, Arg #status predicted
 F:246,268,285-304,328-345,416-429,455-472/Disulfide bonds: #status predicted
 F:287,311/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted
 F:499,520/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted

alignment_scores: Length: 120
 Quality: 135.00 Gaps: 5
 Ratio: 1.667 Percent Identity: 33.333
 Percent Similarity: 67.500

alignment_block:
 US-09-627-165b-13 x S32430 ..

Align seg 1/1 to: S32430 from: 1 to: 527

```

1  GCCAGATTCATCCATCCTGAGAGCTTCGCCGCAATTACAGTGG 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||
165  AAlargyrrArgrrrPheIleSerTrrArgValGlyValSerlleArgrrrAs 181
51  GGAGTCCTCCACCAACATGATGCTCGAGCTGAGAGAGTGG 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||
181  nThrAlrPheGlnProAspAlaAlaMetleuSerleuGlnAsnAsnTPA 198
101  GTGAGCATATCCACCAAGTCCAGCATCCAGATGGCATTTTAAATAC 150
|||||:|||||:|||||:|||||:|||||:|||||:|||||
198  sPAnleuSerGlyGlyValGlnGlnSerValGlnAspThrProAsn 214
151  CAAATAAGATTCAGATTTCCGCCGCTTAC.....TTGTGAGAGTGG 194
|||||:|||||:|||||:|||||:|||||:|||||:|||||
215  AlAlValHrleuArg...SerValAsnAsnGlnProVallleValAspSe 230
195  CAAATGTCGCGAGTATCTCCAGCTTCGCGCATGTTGTCGATCCA 244
|||||:|||||:|||||:|||||:|||||:|||||:|||||
230  rleuThrHlSGlnSerValAlaValleuAlaMetleuPheValCysA 247
245  GTGCTGAGCATTCCTCTCTCTCGACCACTTCGCGCTGCTCTTAAGG 294
|||||:|||||:|||||:|||||:|||||:|||||:|||||
247  sn.....ProProAsnAla.....AsnGlnSerProleuLeuIleArg 259
295  TCCGTCGATGAGTGGCGCAACGATGCTGACCTGCTTTTCCGAAACCCAG 344
|||||:|||||:|||||:|||||:|||||:|||||:|||||
260  SerlleValGlnGlnSerlys...lleCysSerSerArgrrrGlnProHr 275
345  CCGCATC 354
|||||:|||||:|||||:|||||:|||||:|||||:|||||
275  rValAlrArglle 278

```

seq_name: p1r2:S62627

seq_documentation_block:
 agglutinin I precursor - European elder
 C:Species: Sambucus nigra (European elder)
 C:Date: 24-Aug-1996 #sequence_revision 13-Mar-1997 #text_change 20-Aug-1999
 C:Accession: S62627; S62619
 R:van Damme, E.J.M.; Barre, A.; Rouge, P.; van Leuven, F.; Peumans, W.J.
 Eur. J. Biochem. 235, 128-137, 1996
 A:Title: The NeuAc(alpha-2,6)-Gal/GalNAc-binding lectin from elderberry (Sambucus nigra)
 A:Reference numbers: S62619; M01D:96202926
 A:Accession: S62627

A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-570 <VAN>
 A:Cross-references: EMBL:U27122; NID:g1141772; PIDN:AAC49158.1; PID:g1141773
 A:Accession: S62619
 A>Status: preliminary
 A:Molecule type: protein
 A:Residues: 29-39;309-319 <VA2>
 C:Superfamily: ricin: rRNA N-glycosidase homology
 F:37-283/Domain: rRNA N-glycosidase homology <RNG>

alignment_scores:
 Quality: 121.50 Length: 127
 Ratio: 1.519 Gaps: 4
 Percent Similarity: 62.992 Percent Identity: 23.197

alignment_block:
 US-09-627-165B-13 x S62627 ..

Align seg 1/1 to: S62627 from: 1 to: 570

```

1 GCCAGATTCAATCCCATGTCGAGGCTTCGCCGCAATTAACATGCG 50
|||||..... |||..... |||..... |||..... |||.....
201 AATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 217
|||||..... |||..... |||..... |||..... |||.....
51 GGAGTCNTCTCCACCAACATGATGATGATGATGATGATGATGATG 100
|||||..... |||..... |||..... |||..... |||.....
217 AATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 234
|||||..... |||..... |||..... |||..... |||.....
101 GTGCAGATTCACCCAGTCGACGATGACG...GATGCGATTTTAA 147
|||||..... |||..... |||..... |||..... |||.....
234 GATGCGATTTTAAATGATGATGATGATGATGATGATGATGATG 250
|||||..... |||..... |||..... |||..... |||.....
148 ACCCAATAGATGATGATGATGATGATGATGATGATGATGATG 197
|||||..... |||..... |||..... |||..... |||.....
251 GATGCGATTTTAAATGATGATGATGATGATGATGATGATGATG 267
|||||..... |||..... |||..... |||..... |||.....
198 TGTTCGCGATG...ATTCGACCTTGGCATGATGATGATGATG 241
|||||..... |||..... |||..... |||..... |||.....
267 nphatgarguephguleuhtyrilaleuhtyrilaleuhtyril 284
|||||..... |||..... |||..... |||..... |||.....
242 GCAGTGTGCGCATTCCTCTCTCGACGACCCCTCGCGGCTCTCT 291
|||||..... |||..... |||..... |||..... |||.....
284 ysala.....Provalthrserserterysersasnaalale 298
|||||..... |||..... |||..... |||..... |||.....
292 AGTTCGCGATG...GATGCGGCAACGATGATGATGATGATG 323
|||||..... |||..... |||..... |||..... |||.....
299 AlaghtlilelysmetProvalpharglyglylutyrylulys 315
|||||..... |||..... |||..... |||..... |||.....
324 CTGCACTTTTTCGCAACCCGCGGCGCATC 354
|||||..... |||..... |||..... |||..... |||.....
315 lcysserValvalGluValThrArgArgile 325

```

seq_name: p1r2:B39761

seq_documentation_block:
 abrin (clone 7.2) precursor - Indian licorice (fragment)
 N:Contains: rRNA N-glycosidase (EC 3.2.2.22)
 C:Species: Abrus precatorius (Indian licorice)
 C:Date: 21-Feb-1992 #sequence_revision 31-Dec-1993 #text_change 05-Jun-1998
 C:Accession: B39761
 R:Evensen, G.; Mathiesen, A.; Sundan, A.
 J. Biol. Chem. 266, 6848-6852, 1991
 A>Title: Direct molecular cloning and expression of two distinct abrin A-chains.
 A:Reference number: A39761; MUID:91201329
 A:Accession: B39761
 A:Molecule type: DNA
 A:Residues: 1-106 <EVE>
 A:Cross-references: GB:X54873
 C:Superfamily: ricin: rRNA N-glycosidase homology
 C:Keywords: duplication: glycosidase; hydrolase; lectin; toxin
 F:1-106/Product: abrin chain A (fragment) #status predicted <ACH>

F:1-101/Domain: rRNA N-glycosidase homology (fragment) <RNG>
 F:19,22/Active site: Glu, Arg #status predicted

alignment_scores:
 Quality: 99.50 Length: 83
 Ratio: 1.913 Gaps: 1
 Percent Similarity: 62.651 Percent Identity: 33.735

alignment_block:
 US-09-627-165B-13 x B39761 ..

Align seg 1/1 to: B39761 from: 1 to: 106

```

1 GCCAGATTCAATCCCATGTCGAGGCTTCGCCGCAATTAACATGCG 50
|||||..... |||..... |||..... |||..... |||.....
21 AATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 37
|||||..... |||..... |||..... |||..... |||.....
51 GGAGTCNTCTCCACCAACATGATGATGATGATGATGATGATGATG 100
|||||..... |||..... |||..... |||..... |||.....
37 ythralapheginproasproalmetlleserleuhtyrilaleu 54
|||||..... |||..... |||..... |||..... |||.....
101 GTGCAGATTCACCCAGTCGACGATGACG...GATGCGATTTTAA 150
|||||..... |||..... |||..... |||..... |||.....
54 spasnleuserglylaleuhtyrilaleuhtyrilaleuhtyril 70
|||||..... |||..... |||..... |||..... |||.....
151 CAATATGATGATGATGATGATGATGATGATGATGATGATGATG 197
|||||..... |||..... |||..... |||..... |||.....
71 Provalthrserserterysersasnaalaleuhtyrilaleu 87
|||||..... |||..... |||..... |||..... |||.....
198 TGTTCGCGATG...ATTCGACCTTGGCATGATGATGATGATG 246
|||||..... |||..... |||..... |||..... |||.....
87 userhtsProthralalavalleuhtyrilaleuhtyrilaleu 103
|||||..... |||..... |||..... |||..... |||.....

```

seq_name: p1r2:C39761

seq_documentation_block:
 abrin (clone 7.2) precursor - Indian licorice (fragment)
 N:Contains: rRNA N-glycosidase (EC 3.2.2.22)
 C:Species: Abrus precatorius (Indian licorice)
 C:Date: 21-Feb-1992 #sequence_revision 31-Dec-1993 #text_change 20-Aug-1999
 C:Accession: C39761; S14471
 R:Evensen, G.; Mathiesen, A.; Sundan, A.
 J. Biol. Chem. 266, 6848-6852, 1991
 A>Title: Direct molecular cloning and expression of two distinct abrin A-chains.
 A:Reference number: A39761; MUID:91201329
 A:Accession: C39761
 A:Molecule type: DNA
 A:Residues: 1-251 <EVE>
 R:Evensen, G.; Mathiesen, A.; Sundan, A.
 Submitted to the EMBL Data Library, October 1990
 A:Description: Direct molecular cloning of two distinct abrin A-chains.
 A:Reference number: S14471
 A:Accession: S14471
 A:Molecule type: DNA
 A:Residues: 'M',1-251 <EVE>
 A:Cross-references: EMBL:X54872; NID:g16088; PIDN:CAA38654.1; PID:g16089
 C:Superfamily: ricin: rRNA N-glycosidase homology
 C:Keywords: duplication: glycosidase; hydrolase; lectin; toxin
 F:1-251/Product: abrin (clone 7.2) chain A #status predicted <ACH>
 F:7-246/Domain: rRNA N-glycosidase homology <RNG>
 F:74,113,195,196/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted
 F:164,167/Active site: Glu, Arg #status predicted

alignment_scores:
 Quality: 93.50 Length: 83
 Ratio: 1.731 Gaps: 1
 Percent Similarity: 65.060 Percent Identity: 30.120

alignment_block:
 US-09-627-165B-13 x C39761 ..

Align seg 1/1 to: C39761 from: 1 to: 251

```

1 GCCGATTCATCCATCCTGTCGAGCGTTCGCCGCAATTAACAGTGG 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
166 AAlaArgPheIleuArgValGlyValSerLeuGlnAsnTyrTrp 182
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
51 CGAGTCNTCTCCACCAACATGTACATGCTGCAGCTGCAGACGAGTGGG 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
182 YThrAlaPheGlnProAspProAlaMetLeuSerLeuGlnAsnTyr 199
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
101 GTCGACATTCACCCCAAGTCGACGATTCAGGATGCGATTTTATACC 150
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
199 SpAsnLeuSerArgGlyValGlnGlnSerValGlnAspThrPheProAsn 215
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
151 CAATTAAGATG...CAGATTCGCGCGGTAACTTGTCGAGTGAACAA 197
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
216 AAlaValThrLeuArgValAlaAsnGlnProValIleValAspSerLe 232
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
198 TGTGCGCGACGTGATCTCCAGCTTCGCGATCATGTTGTCGATGCACT 246
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
232 UThrIleGlnSerValAlaValIleuAlaLeuMetLeuPheValCysAsn 248
|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

seq_name: p1r1:RLPUGG

seq_documentation_block:

EC 3.2.2.22 alpha-momorcharin precursor [validated] - balsam pear
 N:Alternate names: agglutinin; momordica-
 C:Species: Momordica charantia (balsam pear, bitter melon)
 C:Date: 31-Mar-1992 #sequence revision 31-Mar-1992 #text change 15-Sep-2000
 C:Accession: S14273; A61318; S16490; JN0628; S01670
 R:Ho, W.K.; Liu, S.C.; Shaw, P.C.; Yeung, H.W.; Ng, T.B.; Chan, W.Y.
 Biochim. Biophys. Acta 1088, 311-314, 1991
 A:Title: Cloning of the cDNA of alpha-momorcharin: a ribosome inactivating protein.
 A:Reference number: S14273; MUID:91159486
 A:Accession: S14273
 A:Molecule type: mRNA
 A:Residues: 1-286 <MON>
 A:Cross-references: EMBL:X57682; NID:g19527; PIDN:CAA40869.1; PID:g19528
 R:Li, S.S.L.
 Experientia 36, 524-527, 1980
 A:Title: Purification and partial characterization of two lectins from Momordica charantia
 A:Reference number: A61318; MUID:80201763
 A:Accession: A61318
 A:Molecule type: protein
 A:Residues: 24-50 <LHA>
 A:Note: as a lectin shows agglutinating activity for type-O red blood cells
 R:Montecucchi, P.C.; Lazzerini, A.M.; Barbieri, L.; Stille, F.; Soria, M.; Lappi, D.
 Int. J. Pept. Protein Res. 33, 263-267, 1989
 A:Title: N-terminal sequence of some ribosome-inactivating proteins.
 A:Reference number: S16331; MUID:89326691
 A:Accession: S16490
 A:Molecule type: protein
 A:Residues: 24-68 'X', 70 <MON>
 R:Minami, Y.; Funatsu, G.
 Biosci. Biotechnol. Biochem. 57, 1141-1144, 1993
 A:Title: The complete amino acid sequence of momordin-a, a ribosome-inactivating protein
 A:Reference number: JN0628; MUID:93372485
 A:Accession: JN0628
 A:Molecule type: protein
 A:Residues: 24-107, 'Q', 109-123, 125-147, 'L', 149-154, 'I', 156-205, 'I', 207-208, 'L', 210-214, 'A', Experimental source: seed
 R:Ren, J.; Wang, Y.; Dong, Y.; Stuart, D.I.
 submitted to the Brookhaven Protein Data Bank, January 1994
 A:Reference number: A52272; PDB:1AHC
 A:Contents: annotation: X-ray crystallography, 2.0 angstroms, residues 24-269
 R:Husain, J.; Tickle, I.J.; Wood, S.P.
 submitted to the Brookhaven Protein Data Bank, March 1994
 A:Reference number: A52385; PDB:1JOM
 A:Contents: annotation: X-ray crystallography, 2.16 angstroms, residues 24-86, 'L', 88-269
 R:Huang, Q.; Liu, S.; Tang, Y.; Jin, S.; Wang, Y.
 submitted to the Brookhaven Protein Data Bank, July 1994
 A:Reference number: A67089; PDB:1MRH
 A:Contents: annotation: X-ray crystallography, 2.0 angstroms, residues 24-77, 'R', 79-132,

C:Function: hydrolyzes the N-glycosidic bond of a specific adenosine in 28S rRNA
 A:Description: RNA N-glycosidase; RNA N-glycosidase homology
 C:Superfamily: RNA N-glycosidase; glycosidase; hydrolase; lectin; seed; toxin
 C:Keywords: glycoprotein; glycosidase; hydrolase; lectin; seed; toxin
 F:1-18/Domain: signal sequence #status predicted <PRO>
 F:19-23/Domain: amino-terminal propeptide #status predicted <PRO>
 F:24-269/Product: RNA N-glycosidase homology <RNO>
 F:27-286/Domain: carboxyl-terminal propeptide #status predicted <CTP>
 F:93,183,186/Active site: Tyr, Glu, Arg #status predicted
 F:250/Binding site: carbohydrate (asn) (covalent) #status experimental

alignment_scores:

Quality:	Ratio:	Length:	Gaps:
92.00	1.769	82	3
Percent Similarity: 63.415	Percent Identity: 34.146		

alignment_block:

US-09-627-165B-13 x RLPUGG ..

Align seg 1/1 to: RLPUGG from: 1 to: 286

```

1 GCCGATTCATCCATCCTGTCGAGCGTTCGCCGCAATTAACAGTGG 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
185 AAlaArgPheIleuArgValGlyValSerLeuGlnAsnTyrTrp 201
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
51 CGAGTCNTCTCCACCAACATGTACATGCTGCAGCTGCAGACGAGTGGG 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
201 pGlu.....ValProSerLeuAlaThrIleSerLeuGlnAsnSerTrp 216
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
101 GTCGACATTCACCCCAAGTCGACGATTCAGGATGCGATTTTATACC 144
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
216 eArgLeuSerLysGlnIleGlnLeuAlaGlnGlyAsnAsnLysIlePhe 232
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
145 AATACCAATTAAGATTCGAGATTCGCGCGGTAACTTGTCGACGTCAG 194
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
233 ArgThrProIleValIleValAspAsnLysGlyAsnArgValGlnIleTh 249
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
195 CAATGTT.....CGCAGCTGATCTCCAGCTTCGCGATGCGATGTTG 234
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
249 rAsnValThrSerLysValValThrSerAsnIleGlnLeuLeu 264
|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

seq_name: p1r2:T13535

seq_documentation_block:

Bacillus phage phi-105
 N:acetylmutamoyl-L-alanine amidase homolog
 N:Alternate names: protein 46
 C:Species: Bacillus phage phi-105
 C:Date: 13-Aug-1999 #sequence revision 13-Aug-1999 #text change 20-Jun-2000
 C:Accession: T13535
 R:Kobayashi, K.; Okamura, K.; Inoue, T.; Sato, T.; Kobayashi, Y.
 submitted to the EMBL Data Library, July 1998
 A:Description: Complete nucleotide sequence of Bacillus subtilis phage phi-105.
 A:Reference number: Z17688
 A:Accession: T13535
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-300 <KOB>
 A:Cross-references: EMBL:AB016282; PIDN:BAA36652.1
 C:Superfamily: Bacillus N-acetylmuramoyl-L-alanine amidase

alignment_scores:

Quality:	Ratio:	Length:	Gaps:
87.00	2.071	75	2
Percent Similarity: 56.000	Percent Identity: 33.333		

alignment_block:

US-09-627-165B-13/rev x T13535 ..

Align seg 1/1 to: T13535 from: 1 to: 300

F:10,44,255/Binding site: carbohydrate (asn) (covalent) #status experimental
F:34-258,84-105/Disulfide bonds: #status experimental

alignment_scores:
Quality: 79.00 Length: 59
Ratio: 2.026 Gaps: 1
Percent Similarity: 66.102 Percent Identity: 33.898

alignment_block:
US-09-627-165B-13 x JE0401 ..

Align seg 1/1 to: JE0401 from: 1 to: 261

```
1 GCCAGATTCAATCCCATCCTGAGGCTTCGCCGCAATTACAGTGG 50
|||||
177 AlaArgPheLysTyrIleGluAsnGlnValLysThrAsnPhenArgAs 193
51 GGAGTCNTCTCCACCAACATGTACATGCTGAGCTGAGACGAGTGGG 100
|||
193 PheSer.....ProAsnAspLysValLeuAspLeuGluAsnTrpG 208
101 CTCGACAATCCACCCAGTCCAGCAGTCCAGATGGCATTTTAATACC 150
||::: ||||| :::::|||||::: ||| :::
208 LysLeuSerThrAlaIleHisAsnSerLysAsnGlyAlaLeuProLys 224
151 CAATATAGATTGCAGATTTCGCCCGGT 177
::: |||::: ::: |||
225 ProLeuGluLeuLysAsnAlaAspGly 233
```


OM of: US-09-627-165b-13 to: SwissProt_40:* out_format : pfs
Date: Jul 30, 2002 4:07 PM

About: Results were produced by the Gencore software, version 4.5.
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frame+...n2p.model -DEV=xip
-O=/cgn2.1/USPRO.model/US09627165/runat.30072002_151140_3327/app-query.fasta_1.1012
-DB=SwissProt_40 -OPMT=fastan -SUFFIX=rsip -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPEL=0.000 -LOOPEXT=0.000
-OGAPOP=4.500 -OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=bloms62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pcr
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTPMT=pfs
-NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09627165_ECGN1_1.69 -NCPD=6 -ICPD=3 -LONGLOG
-DEV=US09627165 -WARN_TIMEOUT=30 -NO_XLPHY -WALT -THREADS=1

Search information block:

Query: US-09-627-165b-13
Query length: 357
Database: SwissProt_40:*
Database sequences: 105224
Database length: 38719550
Search time (sec): 36.450000

```
score_list:
Sequence      Std Orig      ZScore      EScore Len      Documentation
SwissProt_40:MLA_VISAL + 317.00 624.82 1.6e-27 254 | P81446 viscum album (europea
SwissProt_40:R1C1_RICCO + 178.00 337.43 7.4e-12 576 | P02879 ricinus communis (cas
SwissProt_40:R1C1_RICCO + 178.00 335.61 9.5e-12 576 | P06750 ricinus communis (cas
SwissProt_40:ABRA_ABRPR + 148.00 277.81 1.7e-08 528 | P11140 abrus precatorius (in
SwissProt_40:ABRC_ABRPR + 138.00 257.10 2.3e-07 562 | P28590 abrus precatorius (in
SwissProt_40:ABRA_ABRPR + 135.00 251.64 4.8e-07 527 | P06077 abrus precatorius (in
SwissProt_40:N1CB_SAMN1 + 108.00 196.66 0.0005 563 | P3183 sambucus nigra (europ
SwissProt_40:R1P1_MOMCH + 92.00 170.60 0.0291 286 | P16094 momordica charantia (
SwissProt_40:R1P1_PHYAM + 84.00 154.24 0.2311 294 | P03464 phytoacca americana
SwissProt_40:R1P2_BRIDI + 83.50 153.61 0.2612 282 | P98184 bryonia dioica (red b
SwissProt_40:R1P5_PHYAM + 79.00 145.25 0.8243 261 | P23339 phytoacca americana
SwissProt_40:R1P5_PHYAM + 76.50 139.29 2.13 289 | P24378 trichosanthos kirilow
SwissProt_40:KARC_ARATH + 74.00 134.64 3.03 267 | P46875 arabidopsis thaliana
SwissProt_40:R1P4_LUCFY + 74.00 134.64 3.03 267 | P00045 luffa cylindrica (smc
SwissProt_40:R1P1_CUCFI + 74.00 134.35 3.05 286 | P09174 cucumis fligere, puta
SwissProt_40:R1P1_CUCFI + 74.00 133.53 3.09 313 | P10297 phytoacca americana
SwissProt_40:EF2_METBU + 74.00 125.83 4.56 730 | O24025 drosophila melanogast
SwissProt_40:SOG_DROME + 73.00 120.61 4.88 1038 | P09989 trichosanthos kirilow
SwissProt_40:EF2_METBU + 72.50 131.23 4.49 289 | P09989 trichosanthos kirilow
SwissProt_40:EF2_METBU + 70.00 117.77 10.00 730 | P09363 mechanococoides methy
SwissProt_40:R1P1_PHYAM + 69.50 113.43 12.09 1054 | P43125 drosophila melanogast
SwissProt_40:R1P1_BRIDI + 69.00 124.15 11.10 290 | P3185 bryonia dioica (red b
SwissProt_40:PEPC_ASPNG + 68.00 118.62 12.29 533 | P33295 aspergillus niger, su
SwissProt_40:CARA_YEAST + 68.50 123.86 12.47 316 | P28495 saccharomyces cerevis
SwissProt_40:R1P3_GELAM + 68.50 122.36 12.82 316 | P33186 gelonium multiflorum
SwissProt_40:ST31_MOUSE + 68.50 111.73 15.57 1018 | P09961 mus musculus (mouse)
SwissProt_40:VIR_FIYPE + 68.00 123.45 14.04 251 | P16089 feline immunodeficienc
SwissProt_40:N1D_FRAL - 68.00 117.69 15.59 473 | P02452 frankia alni, nitroge
SwissProt_40:R1B2_LUCFY + 67.50 122.48 15.96 250 | P22851 luffa cylindrica (smc
SwissProt_40:R1P2_PHYAM + 67.50 108.36 20.66 1182 | P15594 trypanosoma brucei g
SwissProt_40:R1P1_PHYAM + 67.00 120.00 18.66 294 | P56626 trichosanthos angust
SwissProt_40:305B_RAT - 67.00 119.06 18.98 326 | P31210 rattus norvegicus (rat)
SwissProt_40:U125_LITVT + 67.00 113.95 20.84 572 | P29387 infectious iaryngotira
SwissProt_40:PD2_CCNAL + 67.00 110.50 22.19 836 | P06003 candida albicans (yea
SwissProt_40:INTR_BUTFA - 66.50 116.45 22.24 389 | P09052 buto marinus (giant to
SwissProt_40:EF2_HALHA + 66.50 109.74 28.08 728 | P14823 halobacterium halobium
SwissProt_40:FLGI_BUCAI + 66.50 114.84 28.58 372 | P57426 bucheera aphidicola (
SwissProt_40:Y918_HUMAN + 65.50 106.16 33.49 966 | P09499 homo sapiens (human)
SwissProt_40:IR51_MOUSE + 65.50 103.94 34.88 1233 | P35559 mus musculus (mouse)
SwissProt_40:VL3_REOYD + 65.50 103.70 35.04 1267 | P17378 reovirus (type 3 / st
```

```
SwissProt_40:VL3_REOYD + 65.50 103.70 35.04 1267 | P17376 reovirus (type 1 /
SwissProt_40:RT03_MARPO - 65.00 112.51 33.31 430 | P26865 marchantia polymor
SwissProt_40:YD88_SCHPO + 65.00 105.05 38.18 977 | P010495 schizosaccharomyce
SwissProt_40:PKP4_HUMAN + 65.00 103.10 39.57 1211 | P09569 homo sapiens (hum
SwissProt_40:SUBD_BACLI - 64.50 115.61 35.17 274 | P00781 bacillus licheniflo
```

seq_name: SwissProt_40:MLA_VISAL

seq_documentation_block:
ID MLA_VISAL STANDARD; PRT; 254 AA.

```
AC P81446:
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DI 01-MAR-2002 (Rel. 41, Last annotation update)
DE Beta-galactoside specific lectin I A chain (MLA) (MU-I A) (rRNA N-
DE glycosidase) (EC 3.2.2.22).
OS Viscum album (European mistletoe).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Santalales; Viscaceae; Viscum.
OX NCBI_TaxID=3972;
RN [1]
RP SEQUENCE.
RC STRAIN=SUBSP. ALBUM;
RX MEDLINE=97134581; PubMed=9980141;
RA Huguet Soler M., Stoeva S., Schwaborn C., Wilhelm S., Stiefel T.,
RA Voelter W.;
RT FEBS Lett. 399:153-157(1996).
RL
CC -1- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN
CC SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL
CC SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA. THE
CC INO THE CELL OF THE A CHAIN; B CHAINS ARE ALSO RESPONSIBLE FOR
CC CELL AGGUTINATION (LECTIN ACTIVITY).
CC -1- CATALYTIC ACTIVITY: Endohydrolays of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -1- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.
CC -1- MISCELLANEOUS: TWO ISOFORMS OF MLA EXIST: GLYCOSYLATED FORM MLA
CC AND NON-GLYCOSYLATED FORM MLA'.
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY. TYPE 2
CC RIP SUBFAMILY.
DR HSSP: P11140; IABR.
DR InterPro: IPR001574; RIP.
DR Pfam: PF00161; RIP; 1.
DR PRINTS: PR00396; SHIGARICIN.
DR PROSITE: PS00275; SHIGA_RICIN; FALSE_NEG.
KW Hydrolase; Toxin; Repeat; Glycoprotein; Lectin.
FT ACT_SITE 165..165 N-LINKED (GLCNAc...).
FT CARBOHYD 112..112 N-LINKED (GLCNAc...).
FT VARIANT 15 15 E -> D (IN MLA').
FT VARIANT 66 66 V -> I (IN MLA').
FT VARIANT 112 112 N -> T (IN MLA').
FT VARIANT 116 116 P -> T (IN MLA').
FT VARIANT 133 134 DQ -> EE (IN MLA').
FT VARIANT 140 140 T -> S (IN MLA').
FT VARIANT 144 144 F -> Y (IN MLA').
FT VARIANT 151 151 T -> A (IN MLA').
FT VARIANT 179 179 Y -> D (IN MLA').
FT VARIANT 184 184 A -> E (IN MLA').
FT VARIANT 190 190 V -> M (IN MLA').
FT VARIANT 218 218 I -> F (IN MLA').
FT VARIANT 223 224 PP -> ST (IN MLA').
FT VARIANT 231 231 T -> S (IN MLA').
FT VARIANT 235 235 D -> S (IN MLA').
SQ SEQUENCE 254 AA: 28478 MW: 53BAF98D5E0FEE67 CRC64:
```

alignment_scores:
Quality: 317.00 Length: 88
Ratio: 4.342 Caps: 0
Percent Similarity: 82.955 Percent Identity: 71.591

alignment_block:

US-09-627-165b-13 x MLA_VISAL ..

Align seg 1/1 to: MLA_VISAL from: 1 to: 254

```

1 GCCAGATTCATCCATCCATGTCGAGCGTCGCCGCGCAAAATTAACAGTGG 50
  |||||||
167 AAlaArgPheAsnProIleuLeuTrpArgTyrArgGlnTyrIleAsnSerGI 183
51 GGAGCNCNTCTCCAGCAACATGTACATGCTGAGTGGAGAGCAAGTTGGG 100
  |||
183 YAlaSerPheLeuProAspValTyrMetLeuGlnLeuGlnIuThrSerTrpG 200
101 GTGCAAAATTCACCCAGATTCAGCAGATCCAAAGATTCGATTTTAAATACC 150
  ||:|||||
200 YglnglnSerThrIleValGlnHisSerThrAspGlyValPheAsnAsn 216
151 CAAATPAGATTTGCGAGATTTCCGCCGCGTACATTTGTGACAGNTGACCAATGT 200
  |||||||
217 ProIleATgLeuAlaIleProProGlyAsnPhenValThrLeuThrAsnVa 233
201 TGGCAGGTGATCTCCAGCTTGGCGATCATGTTGTTCCAAATGCAAGTGGTC 250
  |||||||
233 lArgspAlIleAlaSerLeuAlaIleMetLeuPhenValCysGlyGlnA 250
251 GGCATTTCTCTCT 264
  ||||
250 rgrProSerSerSer 254

```

seq_name: SwissProt_40:RICI_RICCO

seq_documentation_block:

```

ID RICI_RICCO STANDARD; PRT; 576 AA.
AC P02879; P02880;
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Ricin precursor [Contains: Ricin A chain (rRNA-N-glycosidase)
  (EC 3.2.2.22); Ricin B chain].
OS Ricinus communis (Castor bean).
OC Eukaryotes; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Malpighiales; Euphorbiaceae; Ricinus.
OX NCBI_TaxID=3988;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86067214; PubMed=2999712;
RA Halling K.C., Halling A.C., Murray E.E., Ladin B.F., Houston L.L.,
RA Weaver R.F.;
RT "Genomic cloning and characterization of a ricin gene from Ricinus
  communis."
RL Nucleic Acids Res. 13:8019-8033(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92163016; PubMed=1371405;
RA Tregear J.W., Roberts L.M.;
RT "The lectin gene family of Ricinus communis: cloning of a functional
  ricin gene and three lectin pseudogenes."
RL Plant Mol. Biol. 18:515-525(1992).
RN [3]
RP SEQUENCE OF 12-576 FROM N.A.
RX MEDLINE=85179479; PubMed=3838723;
RA Lamb A., Roberts L.M., Lord J.M.;
RT "Nucleotide sequence of cloned cDNA coding for preproricin."
RL Eur. J. Biochem. 148:265-270(1985).
RN [4]
RP SEQUENCE OF 36-302.
RA Yoshitake S., Funatsu G., Funatsu M.;
RT "Isolation and sequences of peptic peptides, and the complete
  sequence of the chain of ricin-D."
RL Agric. Biol. Chem. 42:1267-1274(1978).
RN [5]
RP SEQUENCE OF 315-576.

```

```

RA Funatsu G., Kimura M., Funatsu M.;
RT "Primary structure of Ala chain of ricin D."
RL Agric. Biol. Chem. 43:2221-2224(1979).
RN [6]
RP CARBOHYDRATE-LINKAGE SITES, AND PARTIAL SEQUENCE.
RX MEDLINE=90344223; PubMed=1368517;
RA Kimura Y., Kusonoki H., Tada M., Takagi S., Funatsu G.;
RT "Structural analyses of sugar chains from ricin A-chain variant."
RL Agric. Biol. Chem. 54:157-162(1990).
RN [7]
RP REVIEW.
RX MEDLINE=21480122; PubMed=11595634;
RA Olmes S., Kozlov J.V.;
RT "Ricin."
RL Toxicol. 39:1723-1728(2001).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
RX MEDLINE=87165983; PubMed=3558397;
RA Monfort W., Villafraña J.E., Monzinger A.F., Ernst S.R., Katzin B.,
RA Rutenber E., Xuong N.H., Hamlin R., Robertus J.D.;
RT "The three-dimensional structure of ricin at 2.8 A."
RL J. Biol. Chem. 262:3598-5403(1987).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF A-CHAIN.
RX MEDLINE=91352004; PubMed=1881881;
RA Katzin B.J., Collins E.J., Robertus J.D.;
RT "Structure of ricin A-chain at 2.5 A."
RL Proteins 10:251-259(1991).
RN [10]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF B-CHAIN.
RX MEDLINE=91352005; PubMed=1881882;
RA Rutenber E., Robertus J.D.;
RT "Structure of ricin B-chain at 2.5-A resolution."
RL Proteins 10:260-269(1991).
RN [11]
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF A-CHAIN.
RX MEDLINE=95082010; PubMed=7990130;
RA Weston S.A., Tucker A.D., Thatcher D.R., Derbyshire D.J.,
RA Pauptit R.A.;
RT "X-ray structure of recombinant ricin A-chain at 1.8-A resolution."
RL J. Mol. Biol. 244:410-422(1994).
RN [12]
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF MUTANT HIS-215.
RX MEDLINE=96374222; PubMed=8780513;
RA Day P.J., Ernst S.R., Frankel A.E., Monzinger A.F., Pascal J.M.,
RA Molina-Svint M.C., Robertus J.D.;
RT "Structure and activity of an active site substitution of ricin A
  chain."
RL Biochemistry 35:11098-11103(1996).
RN [13]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF A-CHAIN.
RX MEDLINE=97240820; PubMed=9086280;
RA Yan X., Hollis T., Svint M., Day P., Monzinger A.F., Milne G.W.,
RA Robertus J.D.;
RT "Structure-based identification of a ricin inhibitor."
RL J. Mol. Biol. 266:1043-1049(1997).
RN [14]
RP MUTAGENESIS.
RX MEDLINE=93165632; PubMed=1287657;
RA Kin Y., Robertus J.D.;
RT "Analysis of several key active site residues of ricin A chain by
  mutagenesis and X-ray crystallography."
RL Protein Eng. 5:775-779(1992).
RN [15]
RP FUNCTION: Ricin is highly toxic to animal cells and to a less
  extent to plant cells. The A chain is responsible for inhibiting
  protein synthesis through the catalytic inactivation of 60S
  ribosomal subunits. It act as a glycosidase that removes that a
  specific adenine residue from an exposed loop of 28S ribosomal
  RNA. As this loop is involved in the binding of elongation
  factors, the modified ribosomes are unable to support protein
  synthesis. The A chain can inactivate a few thousands ribosomes
  per minute, thus inactivating them faster than the cell can make
  new ones. A single A-chain molecule can therefore kill an animal

```

cell. The B chain binds to cell receptors and facilitates the entry into the cell of the A chain; B chains are also responsible for cell agglutination (Lectin activity). It binds to beta-D-galactopyranoside moieties.

-1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.

-1- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.

-1- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).

-1- PTM: THE MAJOR A-CHAIN IS GLYCOSYLATED ONLY IN POSITION 45 AND THE MINOR A-CHAIN IS GLYCOSYLATED ALSO IN POSITION 271.

-1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-INACTIVATING FAMILY. TYPE 2 RIP SUBFAMILY.

-1- SIMILARITY: CONTAINS 2 RICIN B-TYPE LECTIN DOMAINS.

-1- CAUTION: REF.4 AND REF.5 SEQUENCES HAVE A HIGH NUMBER OF CONFLICTS WITH THE SEQUENCE TRANSLATED FROM DNA (REF.1, REF.2 AND REF.3).

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

EMBL: X03179; CAA26939.1; -
 DR EMBL: X52908; CAA37095.1; -
 DR EMBL: X02388; CAA26230.1; -
 DR EMBL: A12892; CAA01058.1; -
 DR FIR: A24041; R1CSD
 DR PDB: 2AAT; 31-JAN-94.
 DR PDB: IAPG; 31-JAN-94.
 DR PDB: IEMP; 31-OCT-93.
 DR PDB: IIFS; 14-JAN-98.
 DR PDB: IIFU; 14-JAN-98.
 DR PDB: IIRC; 31-OCT-93.
 DR PDB: IORS; 16-JUN-97.
 DR PDB: IORT; 16-JUN-97.
 DR PDB: IBR5; 02-SEP-98.
 DR PDB: IBR6; 02-SEP-98.
 DR GLYCOSULEDB: P02879; -
 DR InterPro: IPR001574; RIP.
 DR InterPro: IPR000772; Ricin_B_lectin.
 DR Pfam: PF00652; Ricin_B_lectin; 2.
 DR Pfam: PF00161; RIP; 1.
 DR PRINTS: PR00396; SHIGARICIN.
 DR SMART: SM00458; RICIN_2.
 DR PROSITE: PS50231; RICIN_B_LECTIN; 2.
 DR PROSITE: PS00275; SHIGA_RICIN; 1.
 DR HydroLase: Protein synthesis inhibitor; Toxin; Repeat; Glycoprotein; Lectin; signal; 3D-structure.
 KW SIGNAL; 1
 FT CHAIN 36 302 RICIN A CHAIN.
 FT PEPTIDE 303 314 LINKER PEPTIDE.
 FT CHAIN 315 576 RICIN B CHAIN.
 FT DOMAIN 321 448 RICIN B-TYPE LECTIN 1.
 FT DOMAIN 451 575 RICIN B-TYPE LECTIN 2.
 FT REPEAT 331 373 1-ALPHA.
 FT REPEAT 374 414 1-BETA.
 FT REPEAT 417 449 1-GAMMA.
 FT REPEAT 462 497 2-ALPHA.
 FT REPEAT 501 540 2-BETA.
 FT REPEAT 543 570 2-GAMMA.
 FT ACT_SITE 212 212 INTERCHAIN.
 FT DISULFID 294 318
 FT DISULFID 334 353
 FT DISULFID 377 394
 FT DISULFID 465 478
 FT DISULFID 504 521
 FT CARBOHYD 45 45
 FT CARBOHYD 271 271
 N-LINKED (GLCNAC. . .).
 /FTID-CAR_000080.
 N-LINKED (GLCNAC. . .); IN MINOR A-CHAIN

FT
 FT CARBOHYD 409 409
 FT CARBOHYD 449 449
 FT CONFLICT 76 76
 FT CONFLICT 551 551
 FT STRAND 43 47
 FT TURN 49 50
 FT HELIX 53 67
 FT STRAND 73 74
 FT TURN 75 76
 FT TURN 77 79
 FT TURN 88 90

alignment_scores: Quality: 178.00 Length: 119
 Ratio: 2.342 Gaps: 1
 Percent Similarity: 63.866 Percent Identity: 36.975

alignment_block: US-09-627-165b-13 x RIC1_RICCO ..

Align seg 1/1 to: RIC1_RICCO from: 1 to: 576

1 GCCAGATTCATCCATCTGTCGAGGCGTGGCGCAATTAACAGTGG 50
 214 ALaRphEgInTrIleGluGluMetArGInrArGIlleArGTYrAs 230
 51 GGAGTCNTCTCCACCAACATGTACATGCTCGAGCTGAGACGACTGGG 100
 230 nArGArGSeRrAlaPrOAsPrOSeRrValIleThrLeuGluGuaSenSerTrp 247
 101 GTGCACATTCACCCCAAGTCACGACTGACCAAGAGTGCATTTTATACC 150
 247 lYArGleSeRrThrAlaIleGlnGlnSerAsnGlnAlaIleAlaSer 263
 151 CAATFAGATTGCAGATTTCGCCGCGTACCTTGTGACGNTGAGCAATGT 200
 264 PrOleGlnLeuGlnArGArGAsnGlySerLySPhSeRrValLyTrAsPyA 280
 201 TCGGACGTGATCTCCACGCTTGCGGATGATCATGTTCGAATGCAGTGGC 250
 280 lSerIleuIlePrOleIleAlaIleuMetValTrYArGyAlaIlePrO 297
 251 GGCCATTCCTCCGTCGTGACCCACCTTCGCCGCTGCTCTTAAGTCGCTC 300
 297 rOPrOSeRSeRGlPhSeR.....LeuIleIleArGPrOVal 309
 301 GTGATGGCGCCACGATGTCACTGACCTGACTTTTCGAACCCAGCGGCG 350
 310 ValPrOAsnPhEasnAlaSpValCysMetAsPrOgluPrOleIleValr 326
 351 CAmCGTA 357
 326 gIleVal 328

seq_name: swissprot.40:AGIL_RICCO
 seq_documentation_block: PRT; 564 AA.
 ID AGIL_RICCO STANDARD; PRT; 564 AA.
 AC P06750;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Agglutinin precursor (RCA) [contains: Agglutinin A chain (rRNA N-glycosidase) (EC 3.2.2.22); Agglutinin B chain].
 OS Ricinus communis (Castor bean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot; Rosidae; eurosids I; Malpighiales; Euphorbiaceae; Ricinus.
 OC NCBI_TaxID=3988;
 RN [1]
 RP SEQUENCE FROM N.A.

VARIANT.
 /FTID-CAR_000081.
 N-LINKED (GLCNAC. . .).
 N-LINKED (GLCNAC. . .).
 E -> D (IN REF. 3).
 A -> R (IN REF. 3).

```

FT      CONFLICT      404      404      R -> T (IN REF. 2).
FT      CONFLICT      552      552      F -> V (IN REF. 2).
SQ      SEQUENCE      564 AA: 62851 MW: 145552a72f609759 CRC64;

alignment_scores:
    Quality: 177.00      Length: 119
    Ratio: 2.329      Gaps: 1
    Percent Similarity: 63.866      Percent Identity: 36.975

alignment block:
US-09-627-165B-13 x AGGL_RICCO ..

Align seq 1/1 to: AGGL_RICCO from: 1 to: 564

1 GCCAGATTCAATCCATCCATGNTGTGAGGCTTCGCCGCAATTAACAGTGG 50
|||||.....|||.....:
202 AAtaAgtPhegInIYrIIeGlugIyGlumefAtgThArGArIeArIyAras 218
51 GGAgtCTCTCCACCAAAATCATGTACATGCTGCGAGTGGAGAGGTGGG 100
:::|||||:::.....:
218 nArGArSerAlaProAspProSerValIlehrIeugIusnsrItpG 235
101 GTGCAATCCACCCCAAGTCACAGTCCACAGTCCAGAGATGGATTTCATACC 150
|||||.....:
235 IYArIeUserrThAlaIIeGInglUserrAsngInCIYAlaPheAlaser 251
151 CAATATAGATTGCAGATTTCGCCGCTGACTTGTGACGNTGAGCAATGT 200
|||||.....:
252 PrIoleInleugInArGhrGAsngIySerIySPheAsnValIYrAspA 268
201 TCGGAGAGTGTACTCCAGCTGGCGATCATGTTGTGAAATGAGAGTGTGC 250
|.....:
268 IserIIeUleIleProlleIAlaIeUmeValIYArGyAlaIProP 285
251 GCCATTCTCTCTCTGACACCACTTCGCCGCTGCTGCACAGTCCGTC 300
|||||.....:
285 roPofSerSerInPheSer.....LeuIleIleArGProAl 297
301 GTGATTCGGCGCAACGATGTACCTGCATCTTTTCCGACCCACCGTGG 350
|||||.....:
298 ValProlAsnPhesnAlaAspValCIySmetAspProGluProlleValAr 314
351 CATCGTA 357
|||||
314 gIleVal 316

seq_name: SwissProt_40:ABRA_ABRPR

seq_documentation_block:
ID      ABRA_ABRPR      STANDARD:      PRT:      528 AA.
AC      P11140: P28589;
DT      01-JUL-1989 (Rel. 11, Created)
DT      01-JUN-1994 (Rel. 29, Last sequence update)
DE      01-MAR-2002 (Rel. 41, Last annotation update)
DE      ABrin-a precursor [contains: ABrin-a A chain (rRNA N-glycosidase)
OS      (EC 3.2.2.22); ABrin-a B chain)].
OC      Abus precacorius (Indian Icaricee) (Crab's eye).
OC      Eumetazoa: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
OC      Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidaeae
OC      eucots I; Fabales; Fabaceae; Papilionoideae; Abreae; Abus.
OX      NCBI_TaxId:3816;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=3132798; PubMed=8421313;
RA      Hung C.-H., Lee M.-C., Lee T.-C., Lin J.-Y.;
RT      "Primary structure of three distinct isoabris determined by cDNA
RL      sequencing. Conservation and significance.";
RN      J. Mol. Biol. 229:263-267(1993).
[2]
RP      SEQUENCE OF 1-251.
RC      TISSUE=Seed;

```



```

RT      "The complete amino acid sequence of the A-chain of abrin-a, a toxic
RL      protein from the seeds of Abrus precatorius.",
RN      Agric. Biol. Chem. 52:1095-1097(1988).
RN      [3]
RP      SEQUENCE OF 1-251 FROM N.A.
RC      TISSUE-Leaf;
RX      MEDLINE=91201329; PubMed=2016300;
RA      Evensen G., Mathiesen A., Sundan A.;
RT      "Direct molecular cloning and expression of two distinct abrin
RL      A-chains.";
RN      J. Biol. Chem. 266:6848-6852(1991).
RN      [4]
RP      SEQUENCE OF 262-528.
RX      MEDLINE=92371656; PubMed=1505674;
RA      Chen Y.-L., Chow L.-P., Tsugita A., Lin J.-Y.;
RT      "The complete primary structure of abrin-a B chain.",
RL      FEBS Lett. 309:115-118(1992).
RN      [5]
RP      X-RAY CRYSTALLOGRAPHY (2.14 ANGSTROMS).
RX      MEDLINE=95333188; PubMed=7608980;
RA      Tahirov T.H., Lu T.-H., Liaw Y.-C., Chen Y.-L., Lin J.-Y.;
RT      "Crystal structure of abrin-a at 2.14 A.";
RL      J. Mol. Biol. 250:354-367(1995).
CC      -1- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN
CC      SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL
CC      SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA.
CC      ABRIN-A IS MORE TOXIC THAN RALCIN.
CC      -1- FUNCTION: THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT
CC      FACILITATES THE BINDING OF ABRIN TO THE CELL MEMBRANE THAT
CC      PRECEDES ENDOCYTOSIS.
CC      -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC      specific adenosine on the 28S rRNA.
CC      -1- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.
CC      -1- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN
CC      CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).
CC      -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-
CC      INACTIVATING FAMILY. TYPE 2 RIP SUBFAMILY.
CC      -1- SIMILARITY: CONTRAINS 2 RICIN B-TYPE LECTIN DOMAINS.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on items
CC      use by non-profit institutions as long as their content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sb.ch/announce/
CC      or send an email to license@isb-sb.ch).
CC      -----
DR      EMBL; M98344; AAA32624.1; ALT. INTN'.
DR      EMBL; X54872; -; NOT_ANNOTATED_CDS.
DR      PIR; S32429; TZLSA.
DR      PIR; S24133; S24133.
DR      PDB; 1ABR; 07-FEB-95.
DR      InterPro; IPR001574; RIP.
DR      InterPro; IPR000772; Ricin_B_lectin.
DR      Pfam; PF00652; Ricin_B_lectin; 2.
DR      Pfam; PR00161; RIP; 1.
DR      PRINTS; PR00396; SHIGARICIN.
DR      SMART; SM00458; RICIN; 2.
DR      PROSITE; PS50231; RICIN_B_LECTIN; 2.
DR      PROSITE; PS00275; SHIGA_RICIN; 1.
KM      Hydrolyase; protein synthesis inhibitor; Toxin; Repeat; Glycoprotein;
KM      lectin; 3D-structure.
FT      CHAIN          1          251      ABRIN-A A CHAIN.
FT      PEPTIDE        252      261      LINKER PEPTIDE.
FT      CHAIN          262      528      ABRIN-A B CHAIN.
FT      DOMAIN         273      400      RICIN B-TYPE LECTIN 1.
FT      DOMAIN         403      527      RICIN B-TYPE LECTIN 2.
FT      REPEAT         283      325      1-ALPHA.
FT      REPEAT         326      366      1-BETA.
FT      REPEAT         369      401      1-GAMMA.
FT      REPEAT         414      449      2-ALPHA.
FT      REPEAT         453      492      2-BETA.
FT      REPEAT         495      528      2-GAMMA.

```

FT	ACT SITE	164	164	BY SIMILARITY.
FT	DISULFID	247	269	INTERCHAIN (BY SIMILARITY).
FT	DISULFID	286	305	BY SIMILARITY.
FT	DISULFID	329	346	BY SIMILARITY.
FT	DISULFID	417	430	BY SIMILARITY.
FT	DISULFID	456	473	BY SIMILARITY.
FT	MOD RES	1	1	PROLIDONE CARBOXYLIC ACID.
FT	CARBOHYD	361	361	N-LINKED (GLCNAC. . .).
FT	CARBOHYD	401	401	N-LINKED (GLCNAC. . .).
FT	CONFLICT	202	202	MISSING (IN REF. 2).
FT	CONFLICT	298	298	N -> Y (IN REF. 4).
FT	CONFLICT	427	427	M -> L (IN REF. 4).
FT	CONFLICT	467	467	T -> P (IN REF. 4).
FT	CONFLICT	483	483	V -> L (IN REF. 4).
SO	SEQUENCE	528 AA:	59243 MW:	ALF6BCD5B9A827 CRC64:

alignment_scores:	Quality:	148.00	Length:	129
	Ratio:	1.922	Gaps:	5
Percent Similarity:	59.690	Percent Identity:	33.333	

alignment_block:
US-09-627-165B-13 x ABRA_ABRPR ..

Align seg 1/1 to: ABRA_ABRPR from: 1 to: 528

1	GGCAGATTCATCCCATCTGTCGGAGGCTTCGCGGCNAATTAACAGTGG	50
166	AlaArgPheArgTyrIleSerAsnArgValArgValSerIleGlnThr	182
51	GGAGTCTCTCCACCACAACAGTACATCTCTGAGCTGAGACGAGTGGC	100
182	YthrIaPheGlnProAspAlaAlaMetIleSerLeuGlnAsnTrpA	199
101	GTGCGAATCCACCCAGTCGACGAGCTCCAGAGATTCATTAATACC	150
199	spAsnSerArgGlyValGlnGlnSerValGlnAspThrPheProAsn	215
151	CAATTAAGATTGCAGATTCCGCGCGTAATTGTGACGNTAGGACATGT	200
216	Gln.....ValThrLeuThrAsnI	222
201	TGCGGAC.....GGATCTGCA	217
222	eArgAsnGlnProValIleValAspSerLeuSerHisProThrValAlaVal	239
218	GGTGGCGATCATGTGTTCGATGCAGTGGCGCCATTCCTCTCTC	267
239	AlleuAlaIleuMetLeuPheValCysAsn.....ProProAsnAla...	252
268	GACCAACCTTCGCGCGCTGCTCCTAAGGTCCGTCGTGATGGCGCCACGA	317
253	..AsnGlnSerProLeuLeuIleArgSerIleValGlnIlySerIlys...	267
318	TGTCACCTGCACNTTTCGAGACCCACCTGCGGCATC	354
268	IleCysSerSerArgTyrGlnProThrValArgIle	279

seq_name: Swissprot_40:ABRC_ABRPR

seq_documentation_block:	
ID	ABRC_ABRPR
AC	P28590:
DT	01-DEC-1992 (Rel. 24, Created)
DT	01-DEC-1992 (Rel. 24, Last sequence update)
DT	01-MAR-2002 (Rel. 41, Last annotation update)
DE	Abrin-c precursor [Contains: Abrin-c A chain (rRNA N-glycosylase) (EC 3.2.2.22): Abrin-c B chain)].
DE	Abrus precatorius (Indian licorice) (Crab's eye).
OC	Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae: eustoids I: Fabales: Fabaceae: Papilionoideae: Abreae: Abrus.

seq_name:	Swissprot_40:ABRC_ABRPR
seq_documentation_block:	
ID	ABRC_ABRPR
AC	P28590:
DT	01-DEC-1992 (Rel. 24, Created)
DT	01-DEC-1992 (Rel. 24, Last sequence update)
DT	01-MAR-2002 (Rel. 41, Last annotation update)
DE	Abrin-c precursor [Contains: Abrin-c A chain (rRNA N-glycosylase) (EC 3.2.2.22): Abrin-c B chain)].
DE	Abrus precatorius (Indian licorice) (Crab's eye).
OC	Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae: eustoids I: Fabales: Fabaceae: Papilionoideae: Abreae: Abrus.

OX NCBI_TaxID=3816;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Leaf;
 RX MEDLINE=91266957; PubMed=2050149;
 RA Wood R.A., Lord J.M., Mawiznczak E.J., Platak M.;
 RT "Pleurobabin: genomic cloning, characterisation and the expression of
 RL the A-chain in Escherichia coli."; Eur. J. Biochem. 198;723-732(1991).
 CC -1- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN
 CC SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL
 CC SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA. THE
 CC B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT FACILITATES THE
 CC BINDING OF ABRIN TO THE CELL MEMBRANE THAT PRECEDES ENDOCYTOSIS.
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
 CC specific adenosine on the 28S rRNA.
 CC -1- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.
 CC -1- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN
 CC CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-
 CC INACTIVATING FAMILY. TYPE 2 RIP SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 2 RICIN B-TYPE LECTIN DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X55667; CA39202.1; -
 DR PIR: S16022; S16022.
 DR HSSP: P11140; IABR.
 DR InterPro: IPR001574; RIP.
 DR InterPro: IPR000772; Ricin_B_lectin.
 DR Pfam: PF00652; Ricin_B_lectin; 2.
 DR Pfam: PF00161; RIP; 1.
 DR PRINTS: PRO0396; SHIGARICIN.
 DR SMART: SM00458; RICIN; 2.
 DR PROSITE: PS50231; RICIN_B_LECTIN; 2.
 DR PROSITE: PS00275; SHIGA_RICIN; 1.
 DR HydroLase: Protein synthesis inhibitor; Toxin; Repeat; Glycoprotein;
 KM Lectin; Signal
 FT STGNL 1 34
 FT CHAIN 35 285
 FT PEPTIDE 286 295
 FT CHAIN 296 562
 FT DOMAIN 307 434
 FT DOMAIN 437 561
 FT REPEAT 317 359
 FT REPEAT 360 400
 FT REPEAT 403 435
 FT REPEAT 435 483
 FT REPEAT 487 526
 FT REPEAT 529 562
 FT ACT SITE 198 198
 FT DISULFID 281 303
 FT DISULFID 320 339
 FT DISULFID 363 380
 FT DISULFID 451 464
 FT DISULFID 490 507
 FT MOD_RES 35 35
 FT CARBOHYD 234 234
 FT CARBOHYD 395 395
 FT CARBOHYD 435 435
 SO SEQUENCE 562 AA; 62817 MW; 1FD0ABC7D7BA6278 CRC64;

alignment_scores: Quality: 138.00 Length: 119
 Ratio: 1.792 Gaps: 4

Percent Similarity: 64.706 Percent Identity: 36.134

alignment_block:

US-09-627-165b-13 x ABRRC_ABRPR ..

Align seg 1/1 to: ABRRC_ABRPR from: 1 to: 562

1 GCCAGATTCATCCATCATTGAGAGCTTCCCGCAAAATTAAGTCG 50
 200 Alaagtyrargtyrileserasnargvalglyalserileargthgl 216
 51 GGAGTCNTCCACCAACATGATGTCGAGCTGACGAGTGG 100
 216 ythralahelInProaspproAlameluSerleuglnuasntrpa 233
 101 GTGACATCCACCAAGTCGACGATGCAAGATGTCATTTT...AT 147
 233 spasnleuserglylvalglnInserValglnaspthrPheProasn 249
 148 ACCCAATTAAGATTCGAGATTTCCGCGGTRACTTGTGACGATGACAA 197
 250 AsnvalilleuSerSerileasnarglnProvalValalaspserle 266
 198 TGTTCGCGACGTGATCTGACGCTGCGGATCATGTTGTCGATGCGAGT 247
 266 userHisProthrValalavalaleuAlaleuMetleupheValcysasn. 282
 248 GTGCGCATTTCTCTCTGTCGACCACTTCGCGGCTGCTCTTAAGTCC 297
 283ProProasnaIa.....AsnInserProleuIleuIleargser 295
 298 GTGTCGATGGGCGCACGATGTCACCTGCACCTGACCTTTCCGACCCG 347
 296 IleValgluInserlys...IleCysSerSerargtyrInuProthrva 311
 348 GCGCATC 354
 311 largile 313
 seq_name: SwissProt_40:ABRR_ABRPR
 seq_documentation_block:
 ID ABRRC_ABRPR STANDARD; PRT; 527 AA.
 AC 006077; P81374;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DE 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Abridin-b precursor [contains: Abridin-b A chain (rRNA N-glycosidase)
 OS (EC 3.2.2.22); Abridin-b B chain].
 OS Abridin precursor (Indian licorice) (Crab's eye).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids I; Fabales; Fabaceae; Papilionoideae; Abraceae; Abrus.
 OX NCBI_TaxID=3816;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93132798; PubMed=842133;
 RA Hung C.-H., Lee M.-C., Lee T.-Y.;
 RT "Primary structure of three distinct isoabridins determined by cDNA
 RL sequencing. Conservation and significance."; J. Mol. Biol. 229:263-267(1993).
 RN [2]
 RP SEQUENCE OF 260-527.
 RC TISSUE=Seed;
 RX MEDLINE=93169023; PubMed=7763422;
 RA Kimura M., Sumitawa T., Funatsu G.;
 RT "The complete amino acid sequences of the B-chains of abridin-a and
 RL Biosci. Biotechnol. Biochem. 57:166-169(1993).
 CC -1- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN
 CC SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL
 CC SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA.
 CC ABRIN-A IS MORE TOXIC THAN RICIN.

CC -1- FUNCTION: THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT
 CC FACILITATES THE BINDING OF ABRIN TO THE CELL MEMBRANE THAT
 CC PRECEDES ENDOCYTOSIS.
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
 CC specific adenosine on the 28S rRNA.
 CC -1- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.
 CC -1- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN
 CC CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO THE RIBOSOME-
 CC INACTIVATING FAMILY. TYPE 2 RIP SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 2 RICIN B-TYPE LECTIN DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb-sdb.ch).
 CC -----
 CC EMBL: M98345; AAA32625.1; -.
 CC HSSP: P11140; 1ABR.
 CC InterPro: IPR001574; RIP.
 CC InterPro: IPR000772; Ricin_B_Lectin.
 CC Pfam: PF00652; Ricin_B_Lectin; 2.
 CC Pfam: PF00161; RIP; 1.
 CC PRINTS: PR00396; SHIGARICIN.
 CC SMART: SM00458; RICIN; 2.
 CC PROSITE: PSS0231; RICIN_B_LECTIN; 2.
 CC PROSITE: PSS0275; SHIGA_RICIN; 1.
 CC K01 Hydrolyase; Protein synthesis inhibitor; Toxin; Repeat; Glycoprotein;
 CC Lectin.
 CC KW
 CC FT CHAIN 1 250 ABRIN-B A CHAIN.
 CC FT PEPTIDE 251 260 LINKER PEPTIDE.
 CC FT CHAIN 261 527 ABRIN-B B CHAIN.
 CC FT DOMAIN 272 399 RICIN B-TYPE LECTIN 1.
 CC FT REPEAT 282 324 RICIN B-TYPE LECTIN 2.
 CC FT REPEAT 325 365 1-ALPHA.
 CC FT REPEAT 368 400 1-BETA.
 CC FT REPEAT 413 448 1-GAMMA.
 CC FT REPEAT 452 491 2-ALPHA.
 CC FT REPEAT 494 527 2-BETA.
 CC FT ACT_SITE 163 163 BY SIMILARITY.
 CC FT DISULFID 246 268 INTERCHAIN (BY SIMILARITY).
 CC FT DISULFID 285 304 BY SIMILARITY.
 CC FT DISULFID 328 345 BY SIMILARITY.
 CC FT DISULFID 416 429 BY SIMILARITY.
 CC FT DISULFID 455 472 BY SIMILARITY.
 CC FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID (BY
 CC SIMILARITY).
 CC FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 360 360 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 400 400 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CONFLICT 282 282 N -> D (IN REF. 2).
 CC FT CONFLICT 291 291 D -> N (IN REF. 2).
 CC FT CONFLICT 350 351 AE -> PQ (IN REF. 2).
 CC FT CONFLICT 378 378 S -> N (IN REF. 2).
 CC FT CONFLICT 426 426 L -> M (IN REF. 2).
 CC FT CONFLICT 428 428 Y -> D (IN REF. 2).
 CC FT CONFLICT 431 431 N -> S (IN REF. 2).
 CC FT CONFLICT 484 484 R -> K (IN REF. 2).
 CC FT CONFLICT 491 491 N -> S (IN REF. 2).
 CC FT CONFLICT 493 493 H -> Y (IN REF. 2).
 CC FT CONFLICT 502 502 R -> G (IN REF. 2).
 CC FT CONFLICT 509 509 E -> Q (IN REF. 2).
 CC FT CONFLICT 513 513 H -> W (IN REF. 2).
 CC FT CONFLICT 516 516 H -> T (IN REF. 2).
 CC SEQUENCE 527 AA; 59114 MW; 3253AB490CE9494A CRC64;

alignment_scores: 135.00

Length: 120

Ratio: 1.667 Gaps: 5
 Percent similarity: 67.500 Percent identity: 33.333
 alignment_block:
 US-09-627-165b-13 x ABRB_ABRPR ..
 Align seq 1/1 to: ABRB_ABRPR from: 1 to: 527
 1 GCCAGATTCGATCCCAACNTGTGAGCGCTGCGCCGCAATTAACAGTGG 50
 165 AAlArgrtyrgrphelleSerTyraValGlyValSerIlearghnx 181
 51 GGAGTCNTCTCCACCAACATGTACATGCTCGACGCTGGAGACGATGGG 100
 181 nrfAlaPhglInPrbaspAlaAlaHelIleSerIleGluAsnAsnTTPA 198
 101 CTCGACATCCACCCCAAGCTCCAGACGTCACAGATGGATTTTATATCC 150
 198 spsnLeuSerGlyGlyValGlnGlnSerValGlnAspIhrPheProasn 214
 151 CAATATAGATTGACGATTTCCCGCGGTAC.....TTGTGACGNTGAG 194
 215 AAlvalrhrleuargy...SerValAsnAsnGlnProValIleValaspse 230
 195 CAATGTTGCGAGCTGATCTTCACGCTTGGCGATCATGTGTTGCAATGCA 244
 230 rleuThrHisGlnSerValAlaValLeuAlaLeuMetLeuPheValcysa 247
 245 GTGGTGGCCCATCTCTCTGTCGACACCCCTTCGCGCTGCTCTAAGG 294
 247 sn.....ProPronAsnAla.....AsnGlnSerProLeuLeuIlearg 259
 295 TCCCTGCTGATGCGGCCCAACGATGTCACCTGCATTTTCCGACCCAC 344
 260 SerIleValGlnGlySerIlys...IleCysSerSerargTyrgIuproTh 275
 345 CGTGGCATC 354
 275 rValArgIle 278
 seq_name: SwissProt_40: NIGB_SAMNI
 seq_documentation_block:
 ID NIGB_SAMNI STANDARD; PRT; 563 AA.
 AC P33183; P33184; P93542;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-MAR-2002 (Rel. 41, last sequence update)
 DT 01-MAR-2002 (Rel. 41, last annotation update)
 DE Nigrin b precursor (Agglutinin V) (SNMV) [contains: Nigrin b A chain
 DE (rRNA N-glycosidase) (EC 3.2.2.22); Nigrin b B chain].
 OS Sambucus nigra (European elder).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids II; Adoxaceae; Sambucus.
 ON NCBI_TaxID:4202;
 RX MEDLINE:96215449; PubMed:8647092;
 RP TISSUE-Bark;
 RC TISSUE-Bark;
 RA Van Damme E.J., Barre A., Rouge P., Van Leuven F., Peumans W.J.;
 RT "Characterization and molecular cloning of Sambucus nigra agglutinin V
 RT (nigrin b), a GalNAc-specific type-2 ribosome-inactivating protein
 RT from the bark of elderberry (Sambucus nigra).";
 RL Eur. J. Biochem. 237:505-513(1996).
 RN 121
 RP SEQUENCE OF 26-49 AND 298-321.
 RC TISSUE-Bark;
 RX MEDLINE:94003077; PubMed:8406135;
 RA Gibbs T., Gitoras L., Ferreras J.M., Rojo M.A., Iglesias R.,
 RA Munoz R., Arias F.J., Calonge M., Garcia J.R., Mendez E.;
 RT "Isolation and partial characterization of nigrin b, a non-toxic
 RT novel type 2 ribosome-inactivating protein from the bark of Sambucus
 RT nigra L.";

[illegible]

```

Length: 120
Gaps: 5
cent Identity: 30.000
.
```

```

Align seg 1/1 to: NIGB_SAMNI from: 1 to: 563

1 GCGAGTTTCATCCCATCCTNTGTGGAGGCTTCGCCGCGCAATTAACAATGG 50
190 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 206
51 GGAGTCNCTCCACCAACATGTACATGTCGAGTCGAGAGACGAGTGG 100
206 ATHTSerPheThrProAsnAlaLeuMetLeuSerMetGluAsnAsnTPS 223
101 GTGCGACATCCACCCAGCTCAGCATGTCGCAAGATGCATT ..... 141
223 SerMetSerLeuGlnIleGlnIleGlnAlaGlyAsnAsnValSerProPhe 239
142 TTTAATACCCAAATAGATTGCATATTCGCGCGGTAATTTGTACAGT 191
240 PheGlyThr.....ValGlnLeuLeuAsnTyrAsnHisTh 251
192 GAGCAATGTTCCGAC.....GTGATCTCCACCTTGG 223
251 rHisArgLeuValAspAsnPheGluGlnLeuTyrLysIleThrGlyLea 268
224 CGATATATGTTGATGAGTGCAGTGCAGTGCAGTCTCCCTCTGACCAAC 273
268 IAlLeuLeuPheArgCysSerSer...ProSerAsnAspAsnAlaIle 283
274 CCTTGCCGCTGCTCCTTAAGTCCGTCGTGAGTGGCGCCACAGAT...GT 320
284 ArgMetProLeuAspLeuAlaGlyGluAspAsnLysTyrAsnAspGly 300
321 CACCTGCACCT 330
300 uThrCysThr 303

seq_name: SwissProt_40:RIPL_MOMCH
seq_documentation_block:
ID RIPL_MOMCH STRAND: PRT; 286 AA.
AC P16094; P24697;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Ribosome-inactivating protein momordin I precursor (rRNA
DE 2-glycosylase) (EC 3.2.2.22) (Alpha-momorchatin) (Alpha-MMC).
DE Momordica charantia (Bitter melon) (Balsam pear).
DE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
DE Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
DE eucotsids I; Cucurbitales; Cucurbitaceae; Momordica.
OS NCBI_TaxID=3673;
[1]
RN RNP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RX MEDLINE=91159486; PubMed=2001404;
RA Ho W.K.R., Liu S.C., Shaw P.C., Yeung H.W., Ng T.B., Chan W.Y.;
RT "Cloning of the cDNA of alpha-momorchatin: a ribosome inactivating
RT protein.";
RN RNP Biochim. Biophys. Acta 1088:311-314(1991).
[2]
RN RNP SEQUENCE OF 24-38.
RC TISSUE=Seed;
RX MEDLINE=89326691; PubMed=2753596;
RA Montecuccchi P.C., Lazzarini A.M., Barbieri L., Stilpe F., Soria M.,
RA Lappi D.;
RT "N-terminal sequence of some ribosome-inactivating proteins.";
RN RNP Int. J. Pept. Protein Res. 33:263-267(1989).
[3]
RN RNP SEQUENCE OF 24-70.
RC TISSUE=Seed;
RX MEDLINE=89005108; PubMed=3262509;
RA Casellas P., Dussosoy D., Palasca A.I., Barbieri L.,
RA Guillenot J.C., Ferrara P., Bolognesi A., Gentil P., Stilpe F.;
RT "Trichokinin, a ribosome-inactivating protein from the seeds of
RT Trichostema kirtliowii Maximowicz. Purification, partial

```

RT characterization and use for preparation of immunotoxins.";
 RL Eur. J. Biochem. 176:581-588(1988).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
 RX MEDLINE=94356447; PubMed=8075985;
 RA Ren J., Wang Y., Dong Y., Stuart D.I.;
 RT "The N-glycosidase mechanism of ribosome-inactivating proteins
 implied by crystal structures of alpha-momorcharin.";
 RL Structure 2:7-16(1994).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (2.16 ANGSTROMS).
 RX MEDLINE=94192822; PubMed=8143869;
 RA Husain J., Tickle I.J., Wood S.P.;
 RT "Crystal structure of momordin, a type I ribosome inactivating
 protein from the seeds of Momordica charantia.";
 RL FEBS Lett. 342:154-158(1994).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
 RX MEDLINE=95344383; PubMed=7619070;
 RA Huang O., Liu S., Tang Y., Jin S., Wang Y.;
 RT "Studies on crystal structures, active-centre geometry and
 dehydrating mechanism of two ribosome-inactivating proteins.";
 RL Biochem. J. 309:285-298(1995).
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
 CC specific adenosine on the 28S rRNA.
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY. TYPE 1
 CC RIP SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X57682; CAA40869.1; -
 DR PIR: S14273; RLPUCG.
 DR PIR: S16490; S16490.
 DR PDB: 1AHA; 22-JUN-94.
 DR PDB: 1AHB; 22-JUN-94.
 DR PDB: 1AHC; 22-JUN-94.
 DR PDB: 1MGC; 31-MAY-94.
 DR PDB: 1MRH; 07-FEB-95.
 DR PDB: 1MRI; 07-FEB-95.
 DR GlycoSuiteDB: P16094; -
 DR InterPro: IPR001574; RIP.
 DR Pfam: PF00161; RIP; 1.
 DR PRINTS: PR00396; SHIGARICIN.
 DR PROSITE: PS00275; SHIGA_RICIN; 1.
 DR Protein synthesis inhibitor; Hydrolase; Toxin; signal; Glycoprotein;
 KW 3D-structure.
 FT SIGNAL 1 23
 FT CHAIN 24 269 RIBOSOME-INACTIVATING PROTEIN MOMORDIN I.
 FT PROPEP 270 286 MISSING IN MATURE PROTEIN.
 FT ACT SITE 183 183
 FT CARBOHYD 250 250 N-LINKED (GLCNAC. . .).
 FT /FTID-CAR_000082.
 FT SEQUENCE 286 AA; 31532 MW; E1B013ABEBC21ICF CRC64;
 SQ

alignment_scores:
 Quality: 92.00 Length: 82
 Ratio: 1.769 Gaps: 3
 Percent Similarity: 63.415 Percent Identity: 34.146

alignment_block:

US-09-627-165b-13 x RIP1_MOMCH

Align seg 1/1 to: RIP1_MOMCH from: 1 to: 286

1 GCCAGATTCAATCCATCANTGTGAGCGTTCGCCGCGCAATTAGACGTGG 50

seq_name: SwissProt_40:RIPA_PHYAM
 seq_documentation_block:
 ID RIPA_PHYAM STANDARD; PRT; 294 AA.
 AC 003464; (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 01-OCT-1993 (Rel. 41, Last annotation update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Antiviral protein alpha precursor (PAP-alpha) (Ribosome-inactivating
 DE protein) (rRNA N-glycosidase) (EC 3.2.2.22).
 OS Physalis americana (common pokeweed) (Virginia pokeweed).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Caryophyllidae; Caryophyllales; Physolacaceae; Physolacca.
 OC NCBI_TaxID=3527;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Seed, Leaf, and Root;
 RX MEDLINE=93099240; PubMed=1281438;
 RA Katoaka J., Habuka N., Masuta C., Miyano M., Koizumi A.;
 RT "Isolation and analysis of a genomic clone encoding a pokeweed
 RT antiviral protein.";
 RL Plant Mol. Biol. 20:879-886(1992).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
 RX MEDLINE=95010127; PubMed=7925458;
 RA Ago H., Katoaka J., Tsuge H., Habuka N., Inagaki E., Noma M.,
 RA Miyano M.;
 RT "X-ray structure of a pokeweed antiviral protein, coded by a new
 RT genomic clone, at 0.23-nm resolution. A model structure provides a
 RT suitable electrostatic field for substrate binding.";
 RL Eur. J. Biochem. 225:369-374(1994).
 CC -1- FUNCTION: INHIBITS VIRAL INFECTION OF PLANTS, AND PROTEIN
 CC SYNTHESIS IN VITRO. HAS ALSO BEEN SHOWN TO INHIBIT THE
 CC REPLICATION OF MAMMALIAN VIRUSES. THE PROTEIN MAY PROVIDE A
 CC MEANS OF CELLULAR SUICIDE UPON INVASION BY A VIRUS.
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
 CC specific adenosine on the 28S rRNA.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: Cell wall.
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY. TYPE 1
 CC RIP SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D10600; BAA01451.1; -
 DR PIR: S28421; S28421.
 DR PDB: 1APA; 31-JAN-94.

cores:		Length:	83
Quality:	83.50	Gaps:	4
Ratio:	1.606		

Percent Similarity: 62.651 Percent Identity: 31.325

alignment_block:
US-09-627-165B-13 x R1P2_BRVDI ..

Align seg 1/1 to: R1P2_BRVDI from: 1 to: 282

```

1 GCCAGATTCAATCCCATCCTGAGCGCTTCGCCGCAATTAACAGTGG 50
  |||||
185 AlaArgPheLysTyrIleGluGlnArgValSerGlu.....AsnValG1 199
51 GGAGTCNTCTCCACCAACATGTACATGCTGAGCTGAGACGAGTGGG 100
  |||||
199 yThrLysPheLysProAspProAlaPheLeuSerLeuGlnAsnLatrpg 216
101 GTCCGACATCCACCCAGTCCAG.....CAGTCCAGAGATGCAATTTT 144
  |||||
216 LysLeuSerGluGlnIleGlnIleAlaGlnThrArgGlyGlyGluPhe 232
146 AATACCAATAGATTCAGAT.....TCCGCGGCTAACTTGTGTGAC 188
  |||||
233 AlaArgProValGluIleuArgThrValSerAsnThrProThrPheValTh 249
189 GNTGAGCAATGTCGCGACGTATCTCCAGCTTGGCATCATGTTGTC 237
  |||||
249 rAsnValAsn...SerProValValLysGlyIleAlaLeuLeuLeuTyr 264

```

seq_name: SwissProt_40:R1P2_PHYAM

seq_documentation_block:

```

ID R1P2_PHYAM STANDARD; PRT: 261 AA.
AC P23339;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Activiral protein S (PAP-S) (Ribosome-inactivating protein) (rRNA
  N-glycosidase) (EC 3.2.2.22).
OS Phytolacca americana (Common pokeweed) (Virginian pokeweed);
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllales; Caryophyllales; Phytolaccaceae; Phytolacca.
OX NCBI_TaxID=3527;
RN [1]
RP SEQUENCE.
RC TISSUE=Seed.
RX MEDLINE=91242096; PubMed=1368643;
RA Kung S.S., Kimura M., Funatsu G.;
RT "The complete amino acid sequence of activiral protein from the seeds
  of pokeweed (Phytolacca americana).";
RL Agric. Biol. Chem. 54:3301-3318(1990).
CC -1- FUNCTION: INHIBITS VIRAL INFECTION OF PLANTS, AND PROTEIN
  SYNTHESIS IN VITRO.
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
  specific adenosine on the 28S rRNA.
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY. TYPE 1
  R1P SUBFAMILY.
CC PIR: JEO401; JEO401.
DR HSSP: 003464; JAPA.
DR InterPro: IPR001574; R1P.
DR Pfam: PF00161; R1P.1.
DR PRINTS: PR00396; SHIGARICIN.
DR PROSITE: PS00275; SHIGARICIN.1.
KW Activiral; Protein synthesis inhibitor; Hydrolase; Toxin.
FT ACT_SITE 175 175
FT DISULFID 34 258
FT DISULFID 84 105
FT DISULFID 261 261
FT SEQUENCE 261 AA; 29200 MM; D88B99962FE8399D CRC64;

```

alignment_scores:
Quality: 79.00 Length: 59
Ratio: 2.026 Gaps: 1
Percent Similarity: 66.102 Percent Identity: 33.898

alignment_block:
US-09-627-165B-13 x R1P2_PHYAM ..

Align seg 1/1 to: R1P2_PHYAM from: 1 to: 261

```

1 GCCAGATTCAATCCCATCCTGAGCGCTTCGCCGCAATTAACAGTGG 50
  |||||
177 AlaArgPheLysTyrIleGluAsnGlnValLysThrAsnPhenAsnArgAs 193
51 GGAGTCNTCTCCACCAACATGTACATGCTGAGCTGAGACGAGTGGG 100
  |||||
193 pheSer.....ProAsnAspLysValLeuAspLeuGlnLysntrpg 208
101 GTCCGACATCCACCCAGTCCAGAGTCCAGAGATGCAATTTTAATACC 150
  |||||
208 LysLysIleSerThrAlaIleHisAsnSerLysAsnGlyAlaLeuProLys 224
151 CAATTAAGATTCAGATTCGCCGCGCT 177
  |||||
225 ProLeuGluLeuLysAsnAlaAspGly 233

```

seq_name: SwissProt_40:R1P2_TRIKI

seq_documentation_block:

```

ID R1P2_TRIKI STANDARD; PRT: 289 AA.
AC P24478;
DT 01-MAR-1992 (Rel. 21, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Ribosome-inactivating protein karasurin precursor (rRNA
  N-glycosidase) (EC 3.2.2.22).
OS Trichosanthes kirilowii (Mongolian snake-gourd).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.
OX NCBI_TaxID=3677;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Root tuber.
RX MEDLINE=97356562; PubMed=9212998;
RA Mizukami H., Iida K., Kondo T., Ogihara Y.;
RT "Cloning and bacterial expression of a gene encoding ribosome-
  inactivating proteins, karasurin-A and karasurin-C, from Trichosanthes
  kirilowii var. japonica.";
RL Biol. Pharm. Bull. 20:711-713(1997).
RN [2]
RP SEQUENCE OF 24-270.
RX MEDLINE=92005921; PubMed=1914000;
RA Toyokawa S., Takeda T., Kato Y., Wakabayashi K., Ogihara Y.;
RT "The complete amino acid sequence of an abortifacient protein,
  karasurin.";
RL Chem. Pharm. Bull. 39:1244-1249(1991).
CC -1- FUNCTION: ABORTION-INDUCING PROTEIN. IT INACTIVATES EUKARYOTIC
  60S RIBOSOMAL SUBUNITS.
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
  specific adenosine on the 28S rRNA.
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY. TYPE 1
  R1P SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AB000666; BAA21786.1; -
DR PIR: J00393; J00393.
DR HSSP: P09989; IMR.
DR InterPro: IPR001574; R1P.
DR Pfam: PF00161; R1P.1.

```

```

DR PRINTS: PR00396; SHICARICIN
DR PROSITE: PS00275; SHICARICIN; 1.
KW Antiviral; Protein synthesis inhibitor; Hydrolase; Toxin; Signal.
FT SIGNAL 1 21
FT CHAIN 22 270
FT PROPEP 24 270
FT ACT_SITE 271 289
FT ACT_SITE 183 183
FT SEQUENCE 289 AA; 31704 MW; 883D3E3242887B26 CRC64;
BY SIMILARITY.

alignment_scores:
  Quality: 76.50 Length: 99
  Ratio: 1.195 Gaps: 4
  Percent Similarity: 64.646 Percent Identity: 25.253

alignment_block:
US-09-627-165b-13 x RIPS_TRIKI ..
Align seg 1/1 to: RIPS_TRIKI from: 1 to: 289

1 GCCGATTCATCCATCCATCTGTGGAGGCTTGGCGGCAATTATACAGTGG 50
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
185 Alahrytylylphenileglnlellylsarvalasp.... 199
51 GGAGTCNTCTCCACCAACATGTACATGCTGAGCTGGAGACGAGTGGG 100
  :|||:|||||:|||||:|||||:|||||:|||||:|||||:
200 LysTrpPheLeuProSerLeuAlaIleSerLeuGlnAsnSerTrp 216
101 GTGACATTCACCCACCAAGTCCAG.....CAGTCCAAAGATGCGATT 144
  |||:|||||:|||||:|||||:|||||:|||||:|||||:
216 ErlaleuSerlysglnlelglleAlaSerThrAsnAsnglylnphe 232
145 AATGCCAATAAGATGAGATTTCCGCGGTACTGTTGTGACGAGTGG 194
  :|||:|||||:|||||:|||||:|||||:|||||:
233 GlnThrProValValleuIleAsnAlaGlnAsnGlnArgValThrLeh 249
195 CAATGTCGCGAC.....GTGATCTCCAGCTTGGCGATCATGTTGTCG 238
  :|||:|||||:|||||:|||||:|||||:|||||:
249 rAnValAlaPalaIaGlyValThrSerAsnIleAlaLeuLeu.... 264
239 AATGACGTGTCGCCATTCCTCTCTCCACACCCCTTGGCGCGTGG 285
  :|||:|||||:|||||:|||||:|||||:|||||:
265 ....AsnArgAsnAsnMetAlaAlaIleAspAspValPromeT 278

seq_name: SwissProt_40:KATC_ARATH

seq_documentation_block:
ID KATC_ARATH STANDARD; PRT; 754 AA.
AC P46875;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Kinesin-like protein C.
GN KATC.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN 111
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=94355659; PubMed=8075402;
RA Mitsui H., Nakatani K., Yamaguchi-Shinozaki K., Shinozaki K.,
RA Nishikawa K., Takahashi H.;
RT "Sequencing and characterization of the kinesin-related genes kacb
RT and kact of Arabidopsis thaliana."
RT Plant Mol. Biol. 25:865-876(1994).
CC -1- FUNCTION: POSSIBLE ROLE IN MITOSIS.
CC -1- DOMAIN: COMPOSED OF THREE STRUCTURAL DOMAINS; A SMALL GLOBULAR
CC N-TERMINAL, A CENTRAL ALPHA-HELICAL COILED COIL AND A LARGE
CC GLOBULAR C-TERMINAL WHICH IS RESPONSIBLE FOR THE MOTOR ACTIVITY

```

```

CC (IT HYDROLYZES ATP AND BINDS MICROTUBULES).
CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. NCD
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D21138; BAA04674.1; -.
DR HSSP: P17119; 3KAR.
DR InterPro: IPR001752; kinesin.
DR Pfam: PF00225; kinesin.1.
DR PRINTS: PR00380; KINESINHEAVY.
DR SMART: SM00129; KISC; 1.
DR PROSITE: PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE: PS50067; KINESIN_MOTOR_DOMAIN2; 1.
KW Motor protein; Microtubules; ATP-binding; Coiled coil;
KW Multigene family.
FT DOMAIN 1 45 GLOBULAR.
FT DOMAIN 46 393 COILED COIL.
FT DOMAIN 398 727 KINESIN-MOTOR (BY SIMILARITY).
FT NP_BIND 481 488 ATP (POTENTIAL).
FT SEQUENCE 754 AA; 85030 MW; 76091CD5B5D9C531 CRC64;

alignment_scores:
  Quality: 76.00 Length: 30
  Ratio: 3.167 Gaps: 0
  Percent Similarity: 80.000 Percent Identity: 50.000

alignment_block:
US-09-627-165b-13 x KATC_ARATH ..
Align seg 1/1 to: KATC_ARATH from: 1 to: 754

82 GAGCTGGAGAGAGTGGGTCGACATTCACCCAGTCCAGCTCCAA 131
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
315 GlnIleGlnThrThrCysSerSerGlnSerThrGlnIleArgGlnLeuG 331
132 GGATGCGATTATTAATACCAATTAAGATTCAGATTTC 171
  :|||:|||||:|||||:|||||:|||||:|||||:
331 nAsPArgLeuValaAsnSerGlnArgGlnValSer 344

seq_name: SwissProt_40:RIPA_LUFKY

seq_documentation_block:
ID RIPA_LUFKY STANDARD; PRT; 277 AA.
AC Q00465;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Ribosome-inactivating protein luffin-alpha precursor (rRNA
DE N-glycosidase) (EC 3.2.2.22).
DE Luffa cylindrica (Smooth loofah) (Sponge gourd).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Cucurbitales; Cucurbitaceae; Luffa.
OX NCBI_TaxID=3670;
RN 111
RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RX MEDLINE=92288316; PubMed=1600156;
RA Katoka J., Habuka N., Miyano M., Masuta C., Koizumi A.;
RT "Nucleotide sequence of cDNA encoding alpha-luffin, a ribosome-
RT inactivating protein from luffa cylindrica."
RT Plant Mol. Biol. 18:1199-1202(1992).
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY. TYPE 1

```


RIP SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

CC EMBL: X62371; CAA44229.1; -
 DR PTR: S22494; S22494.
 DR HSSP: P16094; IARC.
 DR InterPro: IPR001574; RIP.
 DR Pfam: PF00161; RIP; 1.
 DR PRINTS: PR00396; SHIGA_RICIN.
 DR PROSITE: PS00275; SHIGA_RICIN; 1.
 KW Antiviral; Protein synthesis inhibitor; Hydrolase; Toxin; Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 277
 FT ACT_SITE 179 179 BY SIMILARITY.
 FT SEQUENCE 277 AA; 30212 MW; EA1FC27998C25AC CRC64;

alignment_scores:
 Quality: 74.00 Length: 80
 Ratio: 1.423 Gaps: 2
 Percent Similarity: 65.000 Percent Identity: 26.250

alignment_block:
 US-09-627-165b-13 x RIP1_LUFCY ..

Align seg 1/1 to: RIP1_LUFCY from: 1 to: 277

```

1 GCCAGATTCAATCCATCMTGTGAGGCTCCGCGCAATTACAGTGG 50
  ::::::::::::::::::::
181 SerArgPheArgTyrIleGluGlnIleIleGluArgIleSer..... 195
51 GGAGTCNTCTCCACCAACATGATGCTGCGAGCTCGGAGAGTGGG 100
  ::::::::::::::::::::
196 LysAsnGlnValProSerLeuAlaThrIleSerLeuGlnAsnGluTrpS 212
101 GTGACATATCCACCAAGTCCAG.....CAGTCCAAAGATGCAATT 144
  ::::::::::::::::::::
212 erAlaIeuSerLysGlnIleGlnIleAlaGlnThrAsnAsnGlyThrPhe 228
145 AATACCAATAAGATTGAGATTCCGCCGGTAACCTTGTGACGNTGG 194
  ::::::::::::::::::::
229 LysThrProValValIleThrAspAspLysGlyGlnArgValGlnIleTh 245
195 CAATGTCGCGAGCTGATCTCCAGCTGGGAGATCATGTTG 234
  ::::::::::::::::::::
245 fAsnValThrSerLysValValThrLysAsnIleGlnIleu 258

```

seq_name: SwissProt_40:RIP1_CUCFI

seq_documentation_block:
 ID RIP1_CUCFI STANDARD; PRT; 286 AA.
 AC Q9FRX4;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Putative ribosome-inactivating protein precursor (rRNA N-glycosidase)
 DE (PC 3.2.2.22).
 OS Cucumis figareli.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Cucurbitales; Cucurbitaceae; Cucumis.
 OX NCBI_TaxID=131071;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yamada T., Ohki S.T., Osaka T.;

RT "Cloning and analysis of a cDNA coding a putative ribosome-
 RT inactivating protein from Cucumis figareli".
 RL Plant Biotechnol. 17:337-340(2000).
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
 CC specific adenosine on the 28S rRNA.
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY. TYPE 1
 CC RIP SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

CC EMBL: AB045560; BAB19677.1; -
 DR InterPro: IPR001574; RIP.
 DR Pfam: PF00161; RIP; 1.
 DR PRINTS: PR00396; SHIGA_RICIN.
 DR PROSITE: PS00275; SHIGA_RICIN; 1.
 KW Antiviral; Protein synthesis inhibitor; Hydrolase; Toxin; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 286
 FT ACT_SITE 185 185 PUTATIVE RIBOSOME-INACTIVATING PROTEIN.
 FT CARBOHYD 103 103 BY SIMILARITY.
 FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT SEQUENCE 286 AA; 31771 MW; 4EPD4966E604DA41 CRC64;

alignment_scores:
 Quality: 74.00 Length: 62
 Ratio: 1.762 Gaps: 1
 Percent Similarity: 67.742 Percent Identity: 29.032

alignment_block:
 US-09-627-165b-13 x RIP1_CUCFI ..

Align seg 1/1 to: RIP1_CUCFI from: 1 to: 286

```

55 TCNTCTCCACCAACATGATGCTGCGAGCTCGGAGAGTGGGGTGG 104
  ::::::::::::::::::::
203 SerLysValProAspLeuAlaIleSerLeuGlnAsnGluTrpSerIe 219
105 ACAATCCACCCAAAGTCCAG.....CAGTCCAAAGATGCAATT 148
  ::::::::::::::::::::
219 uLeuSerLysGlnIleGlnIleAlaLysSerAsnAsnGlyGlnPheGlnT 236
149 CCCAAATAAGATTGAGATTCCGCCGGTAACCTTGTGACGNTGCAAT 198
  ::::::::::::::::::::
236 hrProValLysIleIleAsnAspLysGlyIleLeuThrGluValThrAsn 252
199 GTTCGCGAGCTGATCTCCAGCTGGGAGATCATGTTG 234
  ::::::::::::::::::::
253 ValSerSerLeuValValThrLysAsnIleMetLeu 264

```



```

253 erThAspArgAlaSerSerAspGlnPheIleAspHis.....MetLeu 267
286 CTCCTAAGTCCGTCGTG.....GATGCGGC 311
268 MetIleAlaGProIleLeuValAlaAspValAlaGluValAlaThrAspAlaAs 284
312 CACGATGTCACCTGCACCTTTTCCGAACCCAGCGTCGCATC 354
284 pAsnAspAspThrCysAlaAspProGluProThrValAlaArgIle 298

```

seq_name: sp_plant:Q94BW4

```

seq_documentation_block:
ID   Q94BW4      PRELIMINARY;   PRT;   580 AA.
AC   Q94BW4;
DT   01-DEC-2001 (TREMBLrel. 19, Created)
DT   01-DEC-2001 (TREMBLrel. 19, last sequence update)
DE   TYPE 2 RIBOSOME-INACTIVATING PROTEIN CINNAMOMIN II
DE   PRECURSOR.
OS   Cinnamomum camphora (Camphor tree).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.
OX   NCBI_TaxID=13429;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yang Q., Gong Z.Z., Liu W.Y.;
RT   "Molecular cloning of three type 2 RIP (ribosome-inactivating protein)
RT   genes encoding cinnamomil proteins and study of their expression
RT   patterns."
RL   Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR   EMBL; AY039802; AAK82459.1; -.
KW   Signal.
FT   SIGNAL      1       32      POTENTIAL.
FT   CHAIN       33      580      TYPE 2 RIBOSOME-INACTIVATING PROTEIN
FT   SEQUENCE    580 AA; 64265 MW; 37E4289ECCECBPF CRC64;
SQ

```

alignment_scores:

Quality:	179.00	Length:	131
Ratio:	2.081	Gaps:	6
Percent Similarity:	65.649	Percent Identity:	40.458

alignment_block:

US-09-627-165B-13 x Q94BW4 ..

Align seg 1/1 to: Q94BW4 from: 1 to: 580

```

4 AGATTCATCCCATCNTGTGAGGCTTCGCCGCAATTAACAGTGGGA 53
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
202 ArgPheArgPheIleGluTyrArgValArgGluSerIleThrArgAlaG 218
54 GTCNCTCCACCAACATGTACATGCTGAGCTGAGACGATGGGGTC 103
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
218 UMetPheArgProAspProAlaMetLeuSerLeuGluAsnLysTrpSerA 235
104 GACAAATCCACCAAGTCACAGATCC...AAGATGGCATTTTAAATAC 150
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
235 lAluSerAsnAlaValGlnGlnSerAsnGlnGlyValAlaPheSerSer 251
151 CAATTAAGATTGCAGATTTCGCCGGGTAACCTTTGTGACGNTGACAAATG 200
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
252 ProValGluLeuArgSerIleSerAsnLysProValTyrValGlySerVa 268
201 TCGGGAC...GTGATCTCCAGCTTGGGATGATGTTGTTGATGATGC... 243
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
268 lSerAspArgValIleSerGlyLeuAlaIleMetLeuPheIleCysArgS 285
244 ..AGTGGCCCATTCCTCT.....CTCGACCAACCTTCGCGCGT 285
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
285 erThAspArgAlaSerSerAspGlnPheIleAspHis.....LeuLeu 299

```

```

286 CTCCTAAGTCCGTCGTG.....GATGCGGC 311
300 MetIleAlaGProIleLeuValAlaAspValAlaAspValAlaThrAspAlaAs 316
312 CACGATGTCACCTGCACCTTTTCCGAACCCAGCGTCGCATC 354
316 pAsnAspAspThrCysAlaAspProGluProThrValAlaArgIle 330

```

seq_name: sp_plant:Q94BW3

```

seq_documentation_block:
ID   Q94BW3      PRELIMINARY;   PRT;   580 AA.
AC   Q94BW3;
DT   01-DEC-2001 (TREMBLrel. 19, Created)
DT   01-DEC-2001 (TREMBLrel. 19, last sequence update)
DE   TYPE 2 RIBOSOME-INACTIVATING PROTEIN CINNAMOMIN III
DE   PRECURSOR.
OS   Cinnamomum camphora (Camphor tree).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.
OX   NCBI_TaxID=13429;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yang Q., Gong Z.Z., Liu W.Y.;
RT   "Molecular cloning of three type 2 RIP (ribosome-inactivating protein)
RT   genes encoding cinnamomil proteins and study of their expression
RT   patterns."
RL   Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR   EMBL; AY039803; AAK82460.1; -.
KW   Signal.
FT   SIGNAL      1       32      POTENTIAL.
FT   CHAIN       33      580      TYPE 2 RIBOSOME-INACTIVATING PROTEIN
FT   SEQUENCE    580 AA; 64421 MW; 940D10F01EFB558 CRC64;
SQ

```

alignment_scores:

Quality:	179.00	Length:	131
Ratio:	2.106 <td>Gaps:</td> <td>5 </td>	Gaps:	5
Percent Similarity:	64.885 <td>Percent Identity:</td> <td>38.931</td>	Percent Identity:	38.931

alignment_block:

US-09-627-165B-13 x Q94BW3 ..

Align seg 1/1 to: Q94BW3 from: 1 to: 580

```

4 AGATTCATCCCATCNTGTGAGGCTTCGCCGCAATTAACAGTGGGA 53
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
202 ArgPheArgPheIleGluTyrArgValArgGluSerIleThrArgAlaG 218
54 GTCNCTCCACCAACATGTACATGCTGAGCTGAGACGATGGGGTC 103
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
218 UMetPheArgProAspProAlaMetLeuSerLeuGluAsnLysTrpSerA 235
104 GACAAATCCACCAAGTCACAGATCC...AAGATGGCATTTTAAATAC 150
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
235 lAluSerAsnAlaValGlnGlnSerAsnGlnGlyValAlaPheSerSer 251
151 CAATTAAGATTGCAGATTTCGCCGGGTAACCTTTGTGACGNTGACAAATG 200
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
252 ProValGluLeuArgSerIleSerAsnLysProValTyrValGlySerVa 268
201 TCGGGAC...GTGATCTCCAGCTTGGGATGATGTTGTTGATGATGC... 247
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
268 lSerAspArgValIleSerGlyLeuAlaIleMetLeuPheIleCysArgS 285
248 GTGGCCCATTCCTCT.....CTCGACCAACCTTCGCGCGT 285
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
285 erThAspArgAlaSerSerAspGlnPheIleAspHis.....LeuLeu 299
286 CTCCTAAGTCCGTCGTG.....GATGCGGC 311

```

[illegible]

1 GCCGACATTCGAATCCCATCTGTCGAGCGCTTCCGCCGCAATTAACGACGCG 50
 |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::||
 179 AAlarphgehtnlyrllglucilvglwlketrghrhyrllleargyrhs 195
 51 GGAGTCNTGCCACCAACATGTCATGCTCTGCAGCTGGAGACGACGATTGGG 100
 :|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
 195 nhrgragserAlarproaspProserVallelhrhluGlansSertrpG 212
 101 GTCGACATTCACCCCAAGTCACGACGATCGACGATTCATTTAATAC 150
 |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
 212 lYArgldeuSerThralalieglnluserAanglnlAlalphealaser 228
 151 CAATTAAGATTGACGATTTCCGCCGGTAACTTTGTGACAGTGNAGCAATG 200
 :|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
 229 ProileglnleuglnArgrAanglnSerlySpheSerVallyrAspva 245
 201 TCCGAGCGATGTCGCCAGCTGGCGCATCTATTGTTGCATGACAGTGTC 250
 |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
 245 lserlleuleuleroillelleAlleuueetVallyrArGysAlabrop 262

251 GGCCATTCTCTCTCTGACACCCCTTGGCGCTCTCTTAAGTCGTC 300
 ||| |||:||||| ||| |||:||||| |||
 262 rpfRserSerLtnheser.....LeuLeuIleArgProVal 274
 ||| |||:||||| ||| |||:||||| |||
 301 GTGATGCGCCGACGATGTCACCTGACACTNTTCCGACCCACCGTCGC 350
 ||| |||:||||| ||| |||:||||| |||
 275 ValProAsnPhenAlaAspValcysMetAspProGluProIleValar 291
 ||| |||:||||| ||| |||:||||| |||
 351 CATCGTA 357
 ||| |||:||||| ||| |||:||||| |||
 291 gtlleVal 293

seq_name: sp_plant:Q06076

seq_documentation_block:
 ID Q06076 PRELIMINARY; PRT; 528 AA.
 AC Q06076;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
 DE RNA - GLYCOSIDASE (EC 3.2.2.22) (FRAGMENT)
 OS Abrus precatorius (Indian licorice) (Crab's eye)
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Adreae; Abrus.
 NC NCBITaxID=3816;
 RN |||
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93132798; PubMed=8421313;
 RA Hung C.-H., Lee M.-C., Lee T.-C., Lin J.-Y.;
 RT "Primary structure of three distinct isoenzymes determined by cDNA
 sequencing: conservation and significance";
 RL J. Mol. Biol. 229:263-267(1993).
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
 SPECIFIC ADENOSINE ON THE 28S RNA.
 CC -1- SIMILARITY: TO RIBOSOME INACTIVATING PROTEIN FAMILY.
 CC EMBL: M98346; AAA32626.1; -.
 DR HSRP; P11440; IABR.
 DR InterPro; IPR000772; Ricin_B_lectin.
 DR InterPro; IPR001574; RIP.
 DR Pfam; PF00652; Ricin_B_lectin; 2.
 DR Pfam; PF00161; RIP; 1.
 DR PRINTS; PR00396; SHIGARICIN.
 DR SMART; SM00458; RICIN; 2.
 DR PROSITE; PS50231; RICIN_B_LECTIN; 2.
 DR PROSITE; PS00275; SHIGA_RICIN; UNKNOWN_1.
 KW Hydrolase; Toxin.
 FT NON_TER 1 528
 FT SEQUENCE 528 AA; 58870 MW; 628D42FB8FBE60F8 CRC64;

alignment_scores:
 Quality: 139.00 Length: 119
 Ratio: 1.759 Gaps: 4
 Percent Similarity: 66.387 Percent Identity: 36.134

alignment_block:
 US-09-627-165b-13 x Q06076 ..

Align seg 1/1 to: Q06076 from: 1 to: 528

1 GCCAGATTCATCCCATGTCGAGGCTGCGCCGCAATTAAGTCGTC 50
 ||| |||:||||| ||| |||:||||| |||
 166 AlArGtYrArGcYsIlSeSerAsnArGValGlyValSerIlArGThGl 182
 ||| |||:||||| ||| |||:||||| |||
 51 GGAGTNTCTCCACCAACATGATGATCGAGTCGAGTCGAGAGTCGG 100
 ||| |||:||||| ||| |||:||||| |||
 182 YrTrAlaPhaGlnProAspProAlaMetLeuSerLeuGluAsnArGTrpa 199
 ||| |||:||||| ||| |||:||||| |||
 101 GTGCAATTCACCAAGTCGACGATCC...AAGATGGCATTTTAAAT 147
 ||| |||:||||| ||| |||:||||| |||

199 sPaSnLeuSerGlyValAlGlnGlnSerValGlnAspAlaPhaProAsn 215
 ||| |||:||||| ||| |||:||||| |||
 148 ACCCAATTAAGATTTCAGATTTCGCCGCTGACTTTGACAGNTGACCA 197
 ||| |||:||||| ||| |||:||||| |||
 216 AsnValIlleLeuSerSerIlleAsnArGlnProValValAlaAspSer 232
 ||| |||:||||| ||| |||:||||| |||
 198 TGTTCGCAGCTGATCTCCAGTCAGTCGACATGTTGTCGATGACAGTG 247
 ||| |||:||||| ||| |||:||||| |||
 232 uSerIlSProThrValAlaValLeuAlaLeuMetLeuPheValcysAsn. 248
 ||| |||:||||| ||| |||:||||| |||
 248 GTGGGCATTCCTCTCTCTGACACCCCTTGGCGCTCTCTTAAGTCGTC 297
 ||| |||:||||| ||| |||:||||| |||
 249ProProAsnAla.....AsnGlnSerProLeuIleArgSer 261
 ||| |||:||||| ||| |||:||||| |||
 298 GTGTGATGCGCCGACGATGTCACCTGACACTNTTCCGACCCACCGTC 347
 ||| |||:||||| ||| |||:||||| |||
 262 lIeValGlnGlnSerLys...lIeCysSerSerArGTrGluProThrVa 277
 ||| |||:||||| ||| |||:||||| |||
 348 GCGCATC 354
 ||| |||:||||| ||| |||:||||| |||
 277 lArGtle 279

seq_name: sp_plant:Q9M653

seq_documentation_block:
 ID Q9M653 PRELIMINARY; PRT; 603 AA.
 AC Q9M653;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
 DE RNA - GLYCOSIDASE (EC 3.2.2.22).
 GN R1PT.
 OS Polygonatum multiflorum.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales;
 OC Convallariaceae; Polygonatum.
 NC NCBITaxID=45371;
 RN |||
 RP SEQUENCE FROM N.A.
 RX MEDLINE=2024177; PubMed=10785398;
 RA Van Damme E.J., Hao Q., Charrels D., Barre A., Rouge P., Van Leuven F.,
 RA Pennuans W.J.;
 RT "Characterization and molecular cloning of two different type 2
 RT ribosome-inactivating proteins from the monocotyledonous plant
 RT Polygonatum multiflorum";
 RL Eur. J. Biochem. 267:2746-2759(2000).
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
 SPECIFIC ADENOSINE ON THE 28S RNA.
 CC -1- SIMILARITY: TO RIBOSOME INACTIVATING PROTEIN FAMILY.
 CC EMBL; AF213984; AF37219.1; -.
 DR HSRP; P02879; ZMAT.
 DR InterPro; IPR000772; Ricin_B_lectin.
 DR InterPro; IPR001574; RIP.
 DR Pfam; PF00652; Ricin_B_lectin; 6.
 DR Pfam; PF00161; RIP; 1.
 DR PRINTS; PR00396; SHIGARICIN.
 DR SMART; SM00458; RICIN; 2.
 DR PROSITE; PS50231; RICIN_B_LECTIN; 2.
 DR PROSITE; PS00275; SHIGA_RICIN; UNKNOWN_1.
 KW Hydrolase; Toxin.
 FT SEQUENCE 603 AA; 66733 MW; 88B61C17CA431B1E CRC64;

alignment_scores:
 Quality: 128.50 Length: 140
 Ratio: 1.606 Gaps: 7
 Percent Similarity: 57.143 Percent Identity: 31.429

alignment_block:
 US-09-627-165b-13 x Q9M653 ..

Align seg 1/1 to: Q9M653 from: 1 to: 603

```

1 GCCGATTCATCCATCCTGTCGAGCTTCGCCGCAATTAACAGTGG 50
  |||||
210 AlaArgPheArgHisIleGluIArgValArgArgSerPheThrAlaI 226
  |||||
51 GGATGCTCTCCCA.....CCAAACATGTACATGCTCGAGCTGG 88
  |||||
226 AASPASpAlnLeuThrPheArgProAspGlyLeuMetLeuSerMetG 243
  |||||
89 AGACGAGTGGGTCGACAAATCCCAAGTCGACAGTCC...AAGGAT 135
  |||||
243 LuAsnGluTrpProSerMetSerLeuGluValGlnArgSerIleGluGly 259
  |||||
136 GGCATTTTATACCAATAGATTCAG.....AT 167
  |||||
260 GlyIlePheIleGlyValIleGlnIleAspIleSerPheGlnProle 276
  |||||
168 TTCGCGCGTAACCTTGTGACAGNTGACGATGTCGCGAGCTGATCCGA 217
  |||||
276 ValArgValAspAsnPheAsnThrLeuSerArg.....TyrThrM 289
  |||||
218 GCTTGCGCATCATGTTGTCGATGACAGTGGTGGCA..... 255
  |||||
289 etValAlaLeuLeuLeuPheArgCys...GlyHisProArgAlaThrAla 304
  |||||
256 ...TTCCTCTCTCGACACCCCTCGCGCTCTCTAAGTCCGCT 302
  |||||
305 GlyThrSerSerThrThrProAlaIleAlaGlnIleIleArgMetProva 321
  |||||
303 GGATGGGGC.....AAGCATGTCACCTGCACTGTTT 334
  |||||
321 AspValLeuAlaGlyGluGluIleTyrThrAspGluGluThrCysThrValG 338
  |||||
335 CCGAACCCACCTGCGCATC 354
  |||||
338 LysGluProThrArgArgIle 344
  |||||
seq_name: sp_plant:Q9M654

```

```

KW Hydrolyase; Toxin.
SQ SEQUENCE 604 AA; 67060 MW; 6206C2B25F908D4E CRC64;
Alignment_scores:
  Quality: 128.50 Length: 136
  Ratio: 1.669 Gaps: 6
  Percent Similarity: 56.618 Percent Identity: 30.882

```

alignment_block:
us-09-627-165b-13 x Q9M654 ..

Align seg 1/1 to: Q9M654 from: 1 to: 604

```

1 GCCGATTCATCCATCCTGTCGAGCTTCGCCGCAATTAACAGTGG 50
  |||||
213 AlaArgPheArgHisIleGluIArgValArgArgSerIleAlaAspG 229
  |||||
51 GGATGCTCTCCCAACCAACATGTACATGCTCGAGCTGAGCGAGTGGG 100
  |||||
229 nAspThrPheThrProGlySerLeuMetIleThrMetGluValylsTrps 246
  |||||
101 GTCGCAATCCCAACGACGTCGACAGTCCAAAGAT..... 135
  |||||
246 etIlyMetSerGlnGlnValGluArgSerValAsnAspGlnGlyIlePhe 262
  |||||
136 ...GGCATTTTATACCAATAGATTC.....CA 164
  |||||
263 ThrGlyIlePheThrArgThrValGlnIleuLeuAspAsnLeuGlnTh 279
  |||||
165 GATTTCCGCGCGTAACCTTGTGACGNTGACGATGTCGCGAGTGTCT 214
  |||||
279 rLeuAsnIleAspAsnPh.....AsnAlaLeuSerLeuHisT 292
  |||||
215 CCAAGTTCGGCATCATGTTGTCGATGACAGTGGTGGCATTCCTCT 264
  |||||
292 hrMetLeuAlaIleLeuLeuPheArgCys.....ArgThrThrThrSer 306
  |||||
265 CTCGACCAACCCCTCGCG.....CTGCTCTAAGTCT 296
  |||||
307 SerHisAsnThrLeuProAlaIleSerAsnIleValLeuMetGlyGluAs 323
  |||||
297 CGTGTGATGCGCGGCAACGATGTACCTGCACTGTTTTCGAAACCAACG 346
  |||||
323 pTyValAsp...LysAspAspGluIulysCysThrValGlyGluProThra 339
  |||||
347 TGGCGCATC 354
  |||||
339 rGArgIle 341
  |||||
seq_name: sp_plant:Q9M6E9

```

```

seq_documentation_block:
ID Q9M6E9 PRELIMINARY; PRT; 547 AA.
AC Q9M6E9;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE RNA - GLYCOSIDASE (EC 3.2.2.22).
GN AGC.
OS Abrus precatorius (Indian licorice) (Crab's eye);
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Adreace; Abrus.
OC NCBI_TaxID=3816;
OX [1]
RN SEQUENCE FROM N.A.
RP MEDLINE-20102702; PubMed-10636890;
RX Liu C.L., Tsai C.C., Lin S.C., Wang L.I., Hsu C.I., Hwang M.J.,
RA Lin J.Y.;
RT "Primary Structure and Function Analysis of the Abrus precatorius
  Agglutinin A Chain by Site-directed Mutagenesis: Prolog of Amphiphilic
  alpha-Helix H Impairs Protein Synthesis Inhibitory Activity."

```



```

seq_documentation_block:
ID 022415 PRELIMINARY: PRT: 570 AA.
AC 022415:
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE RNA - GLYCOSIDASE PRECURSOR (EC 3.2.2.22).
GN LEC5NA-1F.
OS Sambucus nigra (European elder).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Adoxaceae; Sambucus.
OC NCBI_TaxID=4202;
OX [1]
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=FRUIT;
RC MEDLINE=98200071; PubMed=9541002;
RA Peumans W.J., Roy S., Baire A., Rouge P., van Leuven F.,
RA van Damme E.J.M.;
RA "Elderberry (Sambucus nigra) contains truncated Neu5Ac(alpha-
RT 2,6)Gal/alpha-binding type 2 ribosome-inactivating proteins.";
RL FEBS Lett. 425:35-39(1998).
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RNA.
CC -1- SIMILARITY: TO RIBOSOME INACTIVATING PROTEIN FAMILY.
DR EMBL: AF012899; AAC49989.1; -.
DR HSSP: P02879; 2NAI.
DR InterPro: IPR000772; R1cln_B_lectln.
DR InterPro: IPR001574; R1P.
DR Pfam: PF00652; R1cln_B_lectln; 2.
DR Pfam: PF00161; R1P; 1.
DR PRINTS: PR00396; SHIGARICIN.
DR SMART: SM00438; RICIN; 2.
DR PROSITE: PS50231; RICIN_B_LECTIN; 2.
DR PROSITE: PS00275; SHIGA_RICIN; 1.
KW Hydrolyase; Signal; Toxin.
KM SIGNAL
FT 1 28 POTENTIAL.
FT CHAIN 29 308 TYPE 2 RIBOSOME-INACTIVATING PROTEIN
FT CHAIN 309 570 SNAIF, A CHAIN.
FT CHAIN TYPE 2 RIBOSOME-INACTIVATING PROTEIN
FT SEQUENCE 570 AA; 62752 MW; 7D799D3A20764BA6 CRC64;
SO

alignment_scores:
Quality: 119.50 Length: 127
Ratio: 1.475 Gaps: 4
Percent Similarity: 63.780 Percent Identity: 25.197

alignment_block:
US-09-627-165B-13 x 022415 ..
Align seg 1/1 to: 022415 from: 1 to: 570
1 GCCAGATTCATCCCATCCTGAGGCTTGGCCGCAATTAACAGTGG 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
201 AlaArgPheArgTyrIleGluLeuArgIleArgThrSerIleThrAsp 217
51 GGAGTCNTCTCCACCAACATGATGATGCTGAGTGGAGAGTGGG 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
217 AserIuPheThrProAspIleuMetLeuSerMetCiluAsnAsnTrps 234
101 GTGCACAAATCCACCAAGTCCAGAGTCCAG...GATGCATTTTAAT 147
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
234 eSerMetSerSerGluIleGlnIleGlnIleGlnProGlyGlyIlePhePro 250
148 ACCCAATTAAGATGACATTTCCCGCGTAACCTTGTGAGCNGAGCAA 197
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
251 GlyValValGlnLeuArgAspGluTrpGlnAsnProIleGluValThrAs 267
198 TGATTCCGACAGTGG...ATCTCCAGCTTGGCGATGTTGTTGCAT 241
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

267 nPheArgArgLeuPheGluLeuThrTyrIleAlaValLeuLeuTyrGlyC 284
242 GCAGTGTGGCCCATCTCTCTCTGACACCACTTGCCTGCTCCTA 291
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
284 yAla.....ProValThrSerAsnSerTyrThrAsnAlaIleAsp 298
292 AGTCGCTGCTGG.....GATCGCGCCACAGATGAC 323
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
299 AlaGlnIleIleLeuMetProValPheArgGlyGlyTyrGluValAsp 315
324 CTGCATCTTTTCCGACCAACCACTGCGCATC 354
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
315 lCySerValValGluValThrArgArgIle 325

seq_name: sp.Plant:004367
seq_documentation_block:
ID 004367 PRELIMINARY: PRT: 563 AA.
AC 004367:
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE RNA - GLYCOSIDASE PRECURSOR (EC 3.2.2.22).
GN Sambucus nigra (European elder).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Adoxaceae; Sambucus.
OC NCBI_TaxID=4202;
OX [1]
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=98112023; PubMed=9450339;
RA Van Damme E.J., Roy S., Baire A., Rouge P., Van Leuven F.,
RA Peumans W.J.;
RA "The major elderberry (Sambucus nigra) fruit protein is a lectin
RT derived from a truncated type 2 ribosome-inactivating protein.";
RL Plant J. 12:1251-1260(1997).
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RNA.
CC -1- SIMILARITY: TO RIBOSOME INACTIVATING PROTEIN FAMILY.
DR EMBL: U76524; AAC15886.1; -.
DR HSSP: P02879; 2NAI.
DR InterPro: IPR000772; R1cln_B_lectln.
DR InterPro: IPR001574; R1P.
DR Pfam: PF00652; R1cln_B_lectln; 2.
DR Pfam: PF00161; R1P; 1.
DR PRINTS: PR00396; SHIGARICIN.
DR SMART: SM00438; RICIN; 2.
DR PROSITE: PS50231; RICIN_B_LECTIN; 2.
DR PROSITE: PS00275; SHIGA_RICIN; 1.
KW Hydrolyase; Signal; Toxin.
KM SIGNAL
FT 1 25 POTENTIAL.
FT CHAIN 26 297 RIBOSOME INACTIVATING PROTEIN, A CHAIN.
FT CHAIN 298 563 RIBOSOME INACTIVATING PROTEIN, B CHAIN.
FT SEQUENCE 563 AA; 62336 MW; 3ED2B6C0B796205 CRC64;
SO

alignment_scores:
Quality: 109.00 Length: 89
Ratio: 1.879 Gaps: 3
Percent Similarity: 65.169 Percent Identity: 32.584

alignment_block:
US-09-627-165B-13 x 004367 ..
Align seg 1/1 to: 004367 from: 1 to: 563
1 GCCAGATTCATCCCATCCTGAGGCTTGGCCGCAATTAACAGTGG 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
190 AlaArgPheArgTyrIleGluGlnIleValArgArgSerIleGlnIleTh 206
51 GGAGTCNTCTCCACCAACATGATGATGCTGAGTGGAGAGTGGG 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
206 lAlaGlyPheThrProAsnAlaIleuMetLeuSerMetCiluAsnAsnTrps 223

```



```

151 GAGGCGGAGATTCATCCATCTTTGGAGGCGCTGCCAATACATTAA 200
    |||||||
165 GluAlaAlaArgPheAsnProIleuTrpArgTyrArgGlnTyrIleAs 181
    |||||||
201 CAGCGGGAGATCTTCCCGACATGTACATCTCGAGCTGGAGACTA 250
    |||||||
181 nSerIylAlaSerPheLeuProAspValTyrMetLeuGlnuencIunhrs 198
251 GTTGGGGCCAAACATCCAGCAGATCCAGCATCTACGATGGGCTTTT 300
    |||||||
198 eTTPGlyGlnGlnSerThrGlnValGlnHisSerThrAspGlyValPhe 214
301 AATAACCCATTTGGTGGGTATATCCACCGGTAACTTCGTGACGTGAG 350
    |||||||
215 AsnAsnProIleArgLeuAlaIleProProGlyAsnPhenValThrLeu 231
351 CAATGTCGCGACGTGATCCGACGCTTAGGATCAGTGTGTGTATGTA 400
    |||||||
231 AsnValArgAspValIleAlaSerLeuAlaIleMetLeuPheValCysG 248
401 GGGACCGACCATCTCTCTCC 420
    ::::::::::
248 LysIuArgProSerSerSer 254

```

seq_name: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:AA25983

seq_documentation_block:

ID AA25983 standard; Protein: 254 AA.

AC AA25983;

DT 18-OCT-1999 (first entry)

DE Mistletoe lectin A1 (variant) protein fragment.

KW Mistletoe; lectin; antitumour; immunostimulant; A-chain; MIA; immunity;
 KW ribozyme 28S subunit; non-cytotoxic; T-cell activation; immune response;
 KW lymphokine-producing macrophage; uncontrolled cell growth; treatment;
 KW cancer; cytotoxicity; antigen; isoform; lectin A1.

OS Viscum album.

PN DE19804210-A1.

PD 12-AUG-1999.

PF 03-FEB-1998; 98DE-1004210.

PR 03-FEB-1998; 98DE-1004210.

XX (BIOS-) BIOSYN ARZNEIMITTEL GMBH.

PI Morris P, Stiefel T, Voelker W, Walters P;

DR WPI; 1999-445335/38.

DR N-PDB; AA209107.

XX Preparation of mistletoe lectins in heterologous systems,
 PT particularly for use as anticancer agents and immunostimulants
 XX
 PS Disclosure: Fig 5B; 78pp; German.

CC This invention describes a novel mistletoe lectin (I) and its fragments
 CC which have antitumour and immunostimulatory activity. The A-chain (MIA)
 CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of
 CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and
 CC lymphokine-producing macrophages, so stimulate immunity. (I) and its
 CC fragments are used to treat uncontrolled cell growth (particularly
 CC cancers) and if they lack cytotoxicity, to increase the strength of the
 CC immune response, particularly to a co-administered antigen
 CC (tumour-associated, bacterial or viral). The method allows production of
 CC mistletoe lectin, and its individual chains, in many different isoforms
 CC and on a large scale, at any time of the year. Recombinant products are

CC free from toxins present in natural mistletoe extracts. This sequence
 CC represents a fragment of a mistletoe lectin A1 protein variant.

XX Sequence 254 AA;

alignment_scores:
 Quality: 621.00 Length: 140
 Ratio: 4.634 Gaps: 0
 Percent Similarity: 95.714 Percent Identity: 85.714

alignment_block:

US-09-627-165B-15 x AA25983 ..

Align seg 1/1 to: AA25983 from: 1 to: 254

```

1 TACACATCTGGAGCGATACGCCGGTTCATAGGACAGATCCCTGTGG 50
    |||
115 TyrProAspLeuGlnuArgTyrAlaGlnHisArgAspGlnIleProLeu 131
51 TATAGAGAACTCATTCATCCGCTCGCGGCTGTGTATCCAGCGCA 100
    |||||||
131 YIleAspGlnLeuIleGlnSerValThrAlaLeuArgPheProGlyIys 148
101 GCACCCGGGCGCCAAAGCTCGTCCCTTATATCCATTCAGATGATCTCC 150
    |||||||
148 eTTPGlyGlnGlnSerThrGlnValGlnHisSerThrAspGlyValPhe 164
151 GAGGCGGAGATTCATCCATCTTTGGAGGCGCTGCCAATACATTAA 200
    |||||||
165 GluAlaAlaArgPheAsnProIleuTrpArgTyrArgGlnTyrIleAs 181
201 CAGCGGGAGATCTTCCCGACATGTACATCTCGAGCTGGAGACTA 250
    |||||||
181 nSerIylAlaSerPheLeuProAspValTyrMetLeuGlnuencIunhrs 198
251 GTTGGGGCCAAACATCCAGCAGATCCAGCATCTACGATGGGCTTTT 300
    |||||||
198 eTTPGlyGlnGlnSerThrGlnValGlnHisSerThrAspGlyValPhe 214
301 AATAACCCATTTGGTGGGTATATCCACCGGTAACTTCGTGACGTGAG 350
    |||||||
215 AsnAsnProIleArgLeuAlaIleProProGlyAsnPhenValThrLeu 231
351 CAATGTCGCGACGTGATCCGACGCTTAGGATCAGTGTGTGTATGTA 400
    |||||||
231 AsnValArgAspValIleAlaSerLeuAlaIleMetLeuPheValCysG 248
401 GGGACCGACCATCTCTCTCC 420
    ::::::::::
248 LysIuArgProSerSerSer 254

```

seq_name: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AA247090

seq_documentation_block:

ID AAB47090 standard; Protein: 254 AA.

AC AAB47090;

DT 16-MAY-2001 (first entry)

DE A-chain isoform for biosynthesis of a Korean mistletoe lectin #1.

KW Isoform: A-chain; B-chain; biosynthesis; lectin; Korean mistletoe;
 KW KML; tumour; KM-110; KML-C; KML-B; KML-III; KML-IV; KML-V;
 KW heparin binding protein.

OS Viscum album coloratum.

PN EPI074560-A2.

XX 07-FEB-2001.

PF 27-JUL-2000; 2000EP-0402168.
 XX 27-JUL-1999; 99KR-0030638.
 XX (MIST-) MISTLE BIOTECH CO LTD.
 PA Kim J, Song S, Suh B, Lee K, Doo M, Kwak J, Song B, Yoon T;
 PI Kang T, Park C;
 XX MPI; 2001-171044/18.
 DR N-PSDB; AAC85472.
 XX Novel lectin proteins isolated from Korean mistletoe, useful for
 PT enhancing immunity and effectuating anti-tumoral activity -
 XX Claim 2; Page 25-26; 62pp; English.
 PS The sequences given in ABA47090-92 are isoforms of an A-chain gene
 CC involved in biosynthesis of lectins isolated from Korean mistletoe.
 CC Korean mistletoe lectins (KML) are useful for enhancing immunity and
 CC for treating tumours. The KML's are isolated from a protein fraction
 CC derived from the leaves, stems and fruits of Korean mistletoe, which
 CC is designated KM-110. One of the isolates, KML-C was shown to be
 CC effective against colon 26-M3.1 carcinoma and L5178Y-M25 lymphoma.
 CC
 XX Sequence 254 AA;

alignment_scores: Length: 140
 Quality: 620.00 Gaps: 0
 Ratio: 4.593 Percent Identity: 87.857
 Percent Similarity: 96.429

alignment_block:
 US-09-627-165B-15 x ABA47090 ..

Align seg 1/1 to: ABA47090 from: 1 to: 254

```

1 TACACAGATCTGACGATACGCCGCTCATAGGACGAGATCCCTGGG 50
115 TTTTAAATGATGATGATGATGATGATGATGATGATGATGATGAT 131
51 TATAGGAGATCATTCATTCATTCATTCATTCATTCATTCATTCATTC 100
111 TATAGGAGATCATTCATTCATTCATTCATTCATTCATTCATTCATTC 148
131 TATAGGAGATCATTCATTCATTCATTCATTCATTCATTCATTCATTC 181
101 GCACCCGCGCCCAAGCTGCTTCCTTAAATCCATTCATTCATTCATTC 150
148 TATAGGAGATCATTCATTCATTCATTCATTCATTCATTCATTCATTC 164
151 GAGGCGCGGAGATTCATTCATTCATTCATTCATTCATTCATTCATTC 200
165 GAGGCGCGGAGATTCATTCATTCATTCATTCATTCATTCATTCATTC 181
201 GAGGCGCGGAGATTCATTCATTCATTCATTCATTCATTCATTCATTC 250
181 TATAGGAGATCATTCATTCATTCATTCATTCATTCATTCATTCATTC 198
251 GTTGGGCGCAACATTCACGACGATTCACGATTCACGATTCACGATTC 300
198 GTTGGGCGCAACATTCACGACGATTCACGATTCACGATTCACGATTC 214
301 AATTAACCATTCGTTGGGTATATTCACCGGTAATTCGTTGGGTATGA 400
215 AATTAACCATTCGTTGGGTATATTCACCGGTAATTCGTTGGGTATGA 231
351 CAATGTCGAGAGTATGCGGATTCAGATCATTCGTTGGGTATGA 400
231 TATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 248
401 GGGACGACGATTCCTCC 420
248 TATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 254

```

seqname: /STDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:AAW64661
 seq_documentation_block:
 ID AAW64661 standard; Protein; 252 AA.
 XX
 AC AAW64661;
 XX
 DT 23-OCT-1998 (first entry)
 XX
 DE Mistletoe rMLA variant protein.

XX Lectin A-chain; mistletoe; rMLA; fusion protein; effector; cytotoxic;
 KW intracellular; processing module; protease recognition; targeting module;
 KW internalisation; treatment; disorder; cell proliferation; activation;
 KW autoimmune disease; allergy; tumour; ricin; translocation; ss.
 XX
 OS Viscum album.

XX Key Location/Qualifiers
 FH 1..252
 FT Protein
 FT /note="partial"

PN W09829540-A2.
 XX
 PD 09-JUL-1998.
 XX
 PF 02-JAN-1998; 98WO-EP00009.
 XX
 PR 02-JAN-1997; 97EP-0100012.
 XX
 PA (BRAI-) BRAIN BIOTECHNOLOGY RES & INFORMATION NE.

XX Eck J, Schmidt A, Zinke H;
 DR MPI; 1998-388122/33.
 DR N-PSDB; AAV51343.

PT Nucleic acid encoding fusion protein containing mistletoe lectin A
 PT chain - useful for treatment of proliferative and autoimmune
 PT diseases, allergies and tumours
 XX
 PS Disclosure: Flg 11a'; 115pp; German.

XX This sequence encodes a variant mistletoe lectin A-chain, rMLA. This
 CC sequence can be used in the construction of a fusion protein which
 CC comprises an effector module that is cytotoxic intracellularly, a
 CC processing module covalently bonded to the effector module and
 CC containing a protease recognition sequence, and a targeting module
 CC covalently bonded to the processing module, able to bind specifically to
 CC the surface of a cell so as to mediate internalisation of the fusion
 CC protein. Such a fusion protein can be used for treating disorders
 CC involving proliferation and/or elevated activation of cells, especially
 CC autoimmune disease, allergy and tumours. The proteins can be administered
 CC at 1 ng to 500 mu g/kg/day, or for ex vivo use at 1 pg to 500 ng/ml.
 CC Fusion proteins can develop toxic activity in a wide range of target
 CC cells. The processing module prevents extracellular dissociation, and
 CC fusion proteins based on mistletoe lectin A-chain are far more active
 CC than those based on ricin and do have the associated problems of
 CC non-specific toxicity. The protein may be expressed in a non-glycosylated
 CC form that does not bind to sugar receptors in the liver, and which has a
 CC long half-life in the blood. Where the mistletoe lectin B-chain is used,
 CC it actively assists in translocation of the ML A-chain from the
 CC endoplasmic reticulum to the cytoplasm.

SO Sequence 252 AA;
 alignment_scores: Length: 138
 Quality: 619.00 Gaps: 0
 Ratio: 4.654 Percent Identity: 86.232
 Percent Similarity: 96.377

alignment_block:
US-09-627-165b-15 x AAW64661 ..

Align seg 1/1 to: AAW64661 from: 1 to: 252

```

1  TACACAGATCTGTGAGCATACCGGTCATATAGGAGACAGATCCCTGTGG 50
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
115  TYPProaspLeucluarGTYrAlaGlyHlaArgAspGlnIleProLeuG1 131
51  TATAGAGAACTCATTCGATCCGCTCGGCGCTCGTATACGAGCGGCA 100
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
131  YlleaspGlnleuIleGlnserValIthrAlaLeuArgPheProGlyLys 148
101  GCACCGGGGCCCAAGCTCGTCCCTTAATTCCTCATTCAGATCTCC 150
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
148  erThrArgThrGlnAlaArgSerIleleuIleleuIleGlnMetIleSer 164
151  GAGCGCGGAGATCAATCCCATCTTTTGGAGGGCTCGCAATACATTAA 200
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
165  GluAlaAlaArgPheAsnProIleleuThrArgAlaArgGlnThrIleAs 181
201  CACCGGGGAGTCATTTCTCCGACATGTACATGCTCGAGCTCGAGACTA 250
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
181  nserGlyAlaSerPheLeuProAspValIYrMetleuGlnleuGlnIle 198
251  GTTGGGGGCCAATCCACGCAAGTCCAGAGTCTACGGAGTGGCTTTT 300
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
198  erTrpGlyGlnGlnserThrGlnValGlnHlaSerThrAspGlyAlaPhe 214
301  AATAACCATTTGCGTGGTATATACACCGGTAACCTGCTGAGCTGTGAG 350
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
215  AsnAsnProIleArgLeuAlaIleProProGlyAsnPheValIThrLeu 231
351  CAATGTCGCGAGTGCATCCGAGTCCATGATGATGATGATGATGATGAT 400
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
231  rAsnValArgAspValIleAlaSerLeuAlaIleMetleuPheValCysG 248
401  GGGACCGACCATCT 414
   ::||| |||||
248  IyGlnArgProSer 252

```

seq_name: /std1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:AAW10022

seq_documentation_block:

ID AAW10022 standard; Protein; 253 AA.

XX AAW10022;

DT 18-DEC-1997 (first entry)

DE Prepro mistletoe lectin A chain.

KM Mistletoe; lectin; cytotoxic; A chain; B chain; dimer.

OS Viscum album.

PN EP751221-A1.

PD 02-JAN-1997.

XX 26-JUN-1995; 95EP-0109949.

PR 26-JUN-1995; 95EP-0109949.

XX (MADU) MADAUUS KOELN AG.

PI Baur A, Eck J, Lentzen H, Zinke H;

DR WPI: 1997-054678/06.

DR N-PSDB; AAT70474.

PT Nucleic acid encoding pre:pro form of mistletoe lectin - for

PT therapeutic or diagnostic use

XX Claim 12; Fig 4A; 30pp; German.

CC Mistletoe lectin is a cytotoxic agent that has been used for tumour

CC therapy. It can be used in immunotoxins and medicaments. Nucleic

CC acid fragments can be used in diagnostic methods. Mistletoe lectin (

XX AAT70473) comprises an A chain (AAT70474) and a B chain (AAT70475).

XX Sequence 253 AA;

alignment_scores: Quality: 619.00 Length: 138

Percent Similarity: 96.377 Percent Identity: 86.232

alignment_block:

US-09-627-165b-15 x AAW10022 ..

Align seg 1/1 to: AAW10022 from: 1 to: 253

```

1  TACACAGATCTGTGAGCATACCGGTCATATAGGAGACAGATCCCTGTGG 50
   ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
116  TYPProaspLeucluarGTYrAlaGlyHlaArgAspGlnIleProLeuG1 132
51  TATAGAGAACTCATTCGATCCGCTCGGCGCTCGTATACGAGCGGCA 100
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
132  YlleaspGlnleuIleGlnserValIthrAlaLeuArgPheProGlyLys 149
101  GCACCGGGGCCCAAGCTCGTCCCTTAATTCCTCATTCAGATCTCC 150
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
149  erThrArgThrGlnAlaArgSerIleleuIleleuIleGlnMetIleSer 165
151  GAGCGCGGAGATCAATCCCATCTTTTGGAGGGCTCGCAATACATTAA 200
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
166  GluAlaAlaArgPheAsnProIleleuThrArgAlaArgGlnThrIleAs 182
201  CACCGGGGAGTCATTTCTCCGACATGTACATGCTCGAGCTGTGAGACTA 250
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
182  nserGlyAlaSerPheLeuProAspValIYrMetleuGlnleuGlnIle 199
251  GTTGGGGGCCAATCCACGCAAGTCCAGAGTCTACGGAGTGGCTTTT 300
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
199  erTrpGlyGlnGlnserThrGlnValGlnHlaSerThrAspGlyAlaPhe 215
301  AATAACCATTTGCGTGGTATATACACCGGTAACCTGCTGAGCTGTGAG 350
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
216  AsnAsnProIleArgLeuAlaIleProProGlyAsnPheValIThrLeu 232
351  CAATGTCGCGAGTGCATCCGAGTCCATGATGATGATGATGATGATGAT 400
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
232  rAsnValArgAspValIleAlaSerLeuAlaIleMetleuPheValCysG 249
401  GGGACCGACCATCT 414
   ::||| |||||
249  IyGlnArgProSer 253

```

seq_name: /std1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:AAW90125

seq_documentation_block:

ID AAW90125 standard; Protein; 253 AA.

XX AAW90125;

DT 30-APR-1999 (first entry)

DE Mistletoe ML A-chain protein.

KM ML; mistletoe; lectin; MLA; A-chain; transgenic plant; glycosylation;

KM dimer; immunotoxin; large-scale production; diagnosis; therapeutic;

XX cancer.

alignment_scores:

Quality: 615.00 Length: 137
 Ratio: 4.659 Gaps: 0
 Percent Similarity: 96.350 Percent Identity: 86.131

alignment_block:

US-09-627-165b-15 x AAW64659 ..

Align seg 1/1 to: AAW64659 from: 1 to: 252

```

1  TACACAGATCTGGAGCGCATACGCCGGTCATAGGACAGATCCCTCTGGG 50
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
116 TytrProaspLeuGlnArgTyrAlaGlyHisArgAspGlnIleProLeuG 132
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
51  TATAGAGAACTCATTCATCCGCTCGGGGGCTGCTTATCCAGCGCGCA 100
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
132 YlleAspGlnLeuIleGlnSerValThrAlaLeuArgPheProGlyGlyS 149
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
101 GCACCCGGGCGCAAGCTCGTCCCTTATATCCATTCATGATGATCTCC 150
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
149 eRTPrGlyGlnGlnAlaArgSerIleLeuIleLeuIleGlnMetIleSer 165
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
151 GAGGCGCGAGATTCATCCATCTTTTGGAGGCGCTCCCAATACATTAA 200
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
166 GlnAlaAlaArgPheAsnProIleLeuTyrArgAlaArgGlnTyrIleAs 182
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
201 CAGCGGGAGATCATTTCTTCCGACATGTACATGCTCGAGCTGGAGACTA 250
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
182 nSerGlyAlaSerPheLeuProAspValTyrMetLeuGlnLeuGlnThrS 199
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
251 GTTGGGGCCACATTCACGCAAGTCCAGCACTGACGATGCGGTTT 300
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
199 eRTPrGlyGlnGlnSerThrGlnValGlnHisSerThrAspGlyValPhe 215
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
301 AANAACCCATTTGCGTGGGTATATCCACCGTAACTTCGAGCGTTGAG 350
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
216 AsnAsnProIleArgLeuAlaIleProProGlyAsnPheValThrLeuTh 232
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
351 CAATGTGCGAGATGATCGCCAGCTTATAGGATCATGTTGTTGATGTA 400
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
232 rAsnValArgAspValIleAlaSerLeuAlaIleMetLeuPheValCysG 249
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
401 GGGACCGACCA 411
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
249 YlleAlaArgPro 252
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

seq_name: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT: AAB47091
 seq_documentation_block:
 ID AAB47091 standard; Protein: 254 AA.

XX AAB47091;
 AC AAB47091;
 DT 16-MAY-2001 (first entry)
 XX
 DE A-chain isoform for biosynthesis of a Korean mistletoe lectin #2.
 XX
 KW Isoform; A-chain; B-chain; biosynthesis; lectin; Korean mistletoe;
 KM KML; tumour; KM-110; KML-C; KMNB; KML-III; KML-III;
 KW heparin binding protein.
 XX
 OS Viscum album coloratum.
 OS
 FH Key Location/Qualifiers
 FT Misc-difference 240 /note= "Encoded by YTA"
 XX
 XX EPI074560-A2.
 PD 07-FEB-2001.
 XX

PF 27-JUL-2000; 2000EP-0402168.

PR 27-JUL-1999; 99KR-0030638.

PA (MIST-) MISTLE BIOTECH CO LTD.

PI Kim J, Song S, Suh B, Lee K, Doo M, Kwak J, Song B, Yoon T;
 Kang T, Park C;

DR WPI: 2001-171044/18.

DR N-PSDB; AAC85473.

PT Novel lectin proteins isolated from Korean mistletoe, useful for
 enhancing immunity and effectuating anti-tumoral activity.

PS Claim 4; Page 26-27; 62pp; English.

XX The sequences given in AAB47090-92 are isoforms of an A-chain gene
 involved in biosynthesis of lectins isolated from Korean mistletoe.
 CC Korean mistletoe lectins (KML) are useful for enhancing immunity and
 CC for treating tumours. The KML's are isolated from a protein fraction
 CC derived from the leaves, stems and fruits of Korean mistletoe, which
 CC is designated KM-110. One of the isolates, KML-C was shown to be
 CC effective against colon 26-M3.1 carcinoma and L5178Y-MU25 lymphoma.
 XX

S0 Sequence 254 AA:

alignment_scores:

Quality: 614.00 Length: 140
 Ratio: 4.582 Gaps: 0
 Percent Similarity: 95.714 Percent Identity: 86.429

alignment_block:

US-09-627-165b-15 x AAB47091 ..

Align seg 1/1 to: AAB47091 from: 1 to: 254

```

1  TACACAGATCTGGAGCGCATACGCCGGTCATAGGACAGATCCCTCTGGG 50
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
115 TytrProaspLeuGlnArgTyrAlaGlyHisArgAspGlnIleProLeuG 131
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
51  TATAGAGAACTCATTCATCCGCTCGGGGGCTGCTTATCCAGCGCGCA 100
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
131 YlleAspGlnLeuIleGlnSerValThrAlaLeuArgPheProGlySerA 148
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
101 GCACCCGGGCGCAAGCTCGTCCCTTATATCCATTCATGATGATCTCC 150
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
148 snThrArgAlaGlnAlaArgSerPheIleIleLeuIleGlnMetIleSer 164
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
151 GAGGCGCGAGATTCATCCATCTTTTGGAGGCGCTCCCAATACATTAA 200
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
165 GlnAlaAlaArgPheAsnProIleLeuTyrArgAlaArgGlnTyrIleSe 181
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
201 CAGCGGGAGATCATTTCTTCCGACATGTACATGCTCGAGCTGGAGACTA 250
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
181 rSerIleYlySerPheLeuProAspThrTyrIleLeuGlnLeuGlnThrS 198
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
251 GTTGGGGCCACATTCACGCAAGTCCAGCACTGACGATGCGGTTT 300
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
198 eRTPrGlyGlnGlnSerThrGlnValGlnHisSerThrAspGlyValPhe 214
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
301 AATAACCATTTGCGTGGGTATATCCACCGGTTAACTTCGTGAGCTTGA 350
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
215 AsnAsnProIleArgLeuThrIleSerThrGlyValPheValThrLeuSe 231
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
351 CAATGTGCGAGATGATCGCCAGCTTATAGGATCATGTTGTTGATGTA 400
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
231 rAsnValArgAspValIleAlaSer***AlaIleMetLeuPheValCysG 248
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
401 GGGACCGACCATCTCTCC 420
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
248 lUnspArgProSerSerSer 254
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```
seq_name: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT: AAB47092
seq_documentation_block:
ID   AAB47092 standard; Protein; 256 AA.
XX
AC   AAB47092;
XX
DT   16-MAY-2001 (first entry)
XX
DE   A-chain isoform for biosynthesis of a Korean mistletoe lectin #3.
XX
KW   Isoform: A-chain; B-chain; biosynthesis; lectin; Korean mistletoe;
KW   KML; tumour; KM-110; KML-C; KMBP; KML-IIU; KML-III;
KW   heparin binding protein.
XX
OS   Viscum album coloratum.
XX
PN   EPI074560-A2.
XX
PD   07-FEB-2001.
XX
PE   27-JUL-2000; 2000EP-0402168.
XX
PR   27-JUL-1999; 99KR-0030638.
XX
PA   (MIST-) MISTLE BIOTECH CO LTD.
XX
PI   Kim J, Song S, Suh B, Lee K, Doo M, Kwak J, Song B, Yoon T;
PI   Kang T, Park C;
XX
DR   WPI: 2001-171044/18.
XX
DR   N-PSDB: AAC85474.
XX
PT   Novel lectin proteins isolated from Korean mistletoe, useful for
XX
PS   enhancing immunity and effectuating anti-tumoral activity -
XX
PS   Claim 6; Page 28; 62pp; English.
XX
CC   The sequences given in AAB47090-92 are isoforms of an A-chain gene
XX
CC   involved in biosynthesis of lectins isolated from Korean mistletoe.
XX
CC   Korean mistletoe lectins (KML) are useful for enhancing immunity and
XX
CC   for treating tumours. The KML's are isolated from a protein fraction
XX
CC   derived from the leaves, stems and fruits of Korean mistletoe, which
XX
CC   is designated KM-110. One of the isolates, KML-C was shown to be
XX
CC   effective against colon 26-M3.1 carcinoma and L5178Y-M25 lymphoma.
XX
SQ   Sequence 256 AA;

alignment_scores:
      Quality: 611.00      Length: 139
      Ratio: 4.628      Gaps: 0
      Percent Similarity: 94.964      Percent Identity: 86.331

alignment_block:
us-09-627-165b-15 x AAB47092 ..

Align seg 1/1 to: AAB47092 from: 1 to: 256

1 TACACAGATCTGAGCGATAGCGCGGTATAGGAGCAACGATCCCTGGG 50
|||||
117 TYRALAAspLeuGlnArgTyrAlaGlyHisArgAspArgIleProLeuG1 133
51 TATAGAGAACTCATTCATCTCGCTCGCGCTTCGTTATCCAGGCGCA 100
|||||
133 yArgGlnProLeuIleArgSerValSerAlaLeuAspTyrProGlyG1S 150
101 GCACCCGGGGCCCAAGCTCGTCCCTTATATCTTATTCAGATGATCC 150
|||||
150 erThrArgAlaGlnAlaSerSerIleIleIleValIleIleGlnMetIleSer 166
151 GAGCGCGCGAGATTCATCCATCTTTGGAGGCGTCGCGCATACATTAA 200
```

```
|||||
167 G1uAlaIleArgPheAsnProIleLeuTrpArgAlaArgGlnTyrIleAs 183
201 CAGCGGGAGCTATTCTTCCGACATATACATGCTCGAGCTGGAGACTA 250
183 nSerGlyValSerTyrLeuProAspValTyrMetLeuGlnLeuGlnAlaS 200
251 GTTGGGGGCAACATCCAGCAAGTCAGCAGCTCTACGAGTGGCGTTT 300
200 erTrpGlyGlnGlnSerThrGlnValGlnGlnSerThrAspGlyValPhe 216
301 AATAACCATTTGGTTGGTATATCCACCGGTACTGCTGACGTGAG 350
217 AsnAsnProIleArgLeuGlyIleSerThrGlyAsnPheValIlePhe 233
351 CAATGTCGAGCAGGATCGCCAGCTTAGCAGCATGTTGTTGATGTA 400
233 rAsnValArgAspValIleAlaSerIleuGlyIleMetValPheValCysa 250
401 GGGACCGACATCTTCC 417
250 rGAspArgSerSerSer 255

seq_name: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT: AA25971
seq_documentation_block:
ID   AA25971 standard; Protein; 255 AA.
XX
AC   AA25971;
XX
DT   18-OCT-1999 (first entry)
XX
DE   Mistletoe lectin A chain MLA consensus protein sequence 1.
XX
KW   Mistletoe; lectin; antitumour; immunostimulant; A-chain; MLA; immunity;
KW   ribozyme 28S subunit; non-cytotoxic; T-cell activation; immune response;
KW   lymphokine-producing macrophage; uncontrolled cell growth; treatment;
KW   cancer; cytotoxicity; antigen; isoform.
XX
OS   Viscum album.
XX
FH   Key Location/Qualifiers
FT   Misc-difference 15 /label= Asp, Glu
FT   FT
FT   Misc-difference 63 /label= Gly, Gln
FT   FT
FT   Misc-difference 66 /label= Ile, Val
FT   FT
FT   Misc-difference 75 /label= Leu, Ala
FT   FT
FT   Misc-difference 107 /label= Asp, Arg or none
FT   FT
FT   Misc-difference 113 /label= Asn, Thr
FT   FT
FT   Misc-difference 117 /label= Pro, Thr
FT   FT
FT   Misc-difference 134 /label= Asp, Glu
FT   FT
FT   Misc-difference 141 /label= Ser, Thr
FT   FT
FT   Misc-difference 145 /label= Phe, Tyr
FT   FT
FT   Misc-difference 152 /label= Thr, Ala
FT   FT
FT   Misc-difference 177 /label= Ala, Tyr
FT   FT
FT   Misc-difference 180 /label= Tyr, Asp
FT   FT
FT   Misc-difference 185 /label= Ala, Glu
FT   FT
FT   Misc-difference 191 /label= Val, Met
FT   FT
FT   Misc-difference 219
```

FT /label= Ile, Phe
FT Misc-difference 224
FT /label= Pro, Ser
FT Misc-difference 225
FT /label= Pro, Thr
FT Misc-difference 232
FT /label= Thr, Ser
FT Misc-difference 236
FT /label= Asp, Ser
XX
XX DE19804210-A1.
XX
XX 12-AUG-1999.
XX
XX 03-FEB-1998; 98DE-1004210.
XX
XX 03-FEB-1998; 98DE-1004210.
XX
XX (BIOS-) BIOSYN ARZNEIMITTEL GMBH.
XX
XX Morris P, Stiefel T, Voelter W, Welters P;
XX WPI; 1999-445335/38.
XX
XX Preparation of mistletoe lectins in heterologous systems,
XX particularly for use as anticancer agents and immunostimulants
XX
XX Claim 2; Page 27; 78pp; German.

This invention describes a novel mistletoe lectin (I) and its fragments which have antitumour and immunostimulatory activity. The A-chain (MLA) of the mistletoe lectin binds to, and inactivates, the 28S subunit of ribosomes. Non-cytotoxic forms of (I) activate T-cell and lymphokine-producing macrophages, so stimulate immunity. (I) and its fragments are used to treat uncontrolled cell growth (particularly cancers) and if they lack cytotoxicity, to increase the strength of the immune response, particularly to a co-administered antigen (tumour-associated, bacterial or viral). The method allows production of the mistletoe lectin, and its individual chains, in many different isoforms and on a large scale, at any time of the year. Recombinant products are free from toxins present in natural mistletoe extracts. This sequence represents a consensus sequence of the mistletoe lectin A chain (MLA) described in the invention.

SO Sequence 255 AA;

alignment_scores:
Quality: 599.00 Length: 140
Ratio: 4.538 Gaps: 0
Percent Similarity: 94.286 Percent Identity: 84.286

alignment_block:

US-09-627-165B-15 x AAY25971 ..

Align seg 1/1 to: AAY25971 from: 1 to: 255

1 TACACAGATCTGAGGATACGGCGATCAGGACCAAGATCCCTGTGG 50
||||:|||||||||||||||||||||||||||||||||||||
116 Tyr***AspLeuGlnArgTyrAlaGlyHisArgAspGlnIleProLeu 132
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
51 TATAGAGAACTCATTCATCCGTCCTCGGCGCTTCGTTACGAGCGGGA 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
132 yIle**GlnLeuIleGlnSerVal**AlaLeuArg***ProGlyGlyS 149
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
101 GCACCGCGGCGCAAGCTGTCCTCCCTATATATCCCTCATTCAGATGCC 150
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
149 erThrArg***GlnAlaArgSerIleLeuIleLeuIleGlnMetIleSer 165
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
151 GAGCGCGGAGATTCATCCATCTTTGGAGGCGTCGCCATTAATTA 200
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
166 GluAlaAlaArgPheAsnProIleLeuTyrPArg**ArgGln**IleAs 182

201 CAGCGGAGATCATTTCTCCCGACATGATACATGCTGACCTGAGACTA 250
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
182 nserGly**SerPheLeuProAsp***TyrMetLeuGlnLeuGlnThrS 199
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
251 GTTGGGCGCCACAAATCCAGCAAGTCAGAGCTACGAGTGGCGTTT 300
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
199 erTyrGlyGlnGlnSerThrGlnValGlnHisSerThrAspGlyValPhe 215
301 AATAACCATTTGCGTTGGGTATATCCACGCGTACTTCGTGACGTTGAG 350
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
216 AsnAsnPro**ArgLeuAlaIle****GlyAsnPheValThrLeu** 232
351 CAATGTTCCGCGACGTGATCCGACGCTTAGCGATCATGTTGTTATGTA 400
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
232 *AsnValArg***ValIleAlaSerLeuAlaIleMetLeuPheValCysG 249
401 GCGACCGACCATCTCTCCTCC 420
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
249 LysIuArgProSerSerSer 255

seq_name: /SIDSI/gcgdata/hold-geneseq/geneseq_emb1/AA1999.DAT:AA25974

seq_documentation_block:

ID AAY25974 standard; protein: 255 AA.

AC AAY25974;

DT 18-OCT-1999 (first entry)

DE Mistletoe lectin A chain MLA consensus protein sequence 2.

KW rlyozyme 28S subunit; non-cytotoxic; T-cell activation; immune response;
KW lymphokine-producing macrophage; uncontrolled cell growth; treatment;
KW cancer; cytotoxicity; antigen; isoform.

XX Viscum album.

OS Key Location/Qualifiers

FT Misc-difference 15 /label= Asp, Glu
FT FT
FT Misc-difference 63 /label= Gly, Gln
FT FT
FT Misc-difference 66 /label= Ile, Val
FT FT
FT Misc-difference 75 /label= Leu, Ala
FT FT
FT Misc-difference 107 /label= Asp, Arg, none
FT FT
FT Misc-difference 113 /label= Asn, Thr
FT FT
FT Misc-difference 117 /label= Pro, Thr
FT FT
FT Misc-difference 134 /label= Asp, Glu
FT FT
FT Misc-difference 141 /label= Ser, Thr
FT FT
FT Misc-difference 145 /label= Phe, Tyr
FT FT
FT Misc-difference 152 /label= Thr, Ala
FT FT
FT Misc-difference 177 /label= Ala, Tyr
FT FT
FT Misc-difference 180 /label= Tyr, Asp
FT FT
FT Misc-difference 185 /label= Ala, Glu
FT FT
FT Misc-difference 191 /label= Val, Met
FT FT
FT Misc-difference 219 /label= Ile, Phe
FT FT
FT Misc-difference 224 /label= Phe, Ser

```

251 GTTGGGGCCACATCATCCAGCAAGTCCAGACAGTCTACGATGGCGTTTT 300
      |||||
199 eetrpIyGInGInSerThrGInValGInHisSerThrAspGlyValPhe 215
      |||||
301 AATAACCATTTGGTGGGTATATCCACGGTAACTTCGTGACGTTGAG 350
      |||||
216 AsnAspPro***ArgLeuAlaLeu*****QlyAsnPheValThrLeu** 232
      |||||
351 CAATGTTCGGCAGCTGATGCCAGCTTAGCCATCATCTGTTGATATGA 400
      :|||
232 *AsnValArg**ValIleLeuSerLeuAlaIleMetLeuPheValCysG 249
      |||||
401 GGGACGACCATCTCTCC 420
      :|||
249 LysIuArgProSerSerSer 255

```

251	GTTCGGGGCCACATCCACGCAAGTCCAGACGTCCTGAGATGGCGTTT	300
199		215
301		350
216	AsnAspPro***ArgLeuAlaIle*****GlyAsnPheValThrLeu**	232
351	CAATGTCGCGACGTGATCGCCAGCTTACGATCACTGTTCTTGTATGTA	400
232	*AsnValArg***ValIleAlaSerLeuAlaIleMetLeuPheValCysG	249
401	GGGACCGACCATCTTCCTCC	420
249	LYGUAHqPProSerSer	255
seq_name: /SIDS1/gcgdata/hold-geneseq/geneseq-embL/AA1999.DAT:AAV25977		
seq_documentation_block:		
ID	AAV25977	strand; protein; 255 AA.
XX	AAV25977;	
XX	18-OCT-1999	(first entry)
DE	Mistletoe lectin A chain MLA consensus protein sequence 3.	
XX	Mistletoe lectin; antitumor; immunostimulant; A-chain; MLA; immunity;	
KW	ribosome 26S subunit; non-cytotoxic; T-cell activation; immune response	
KW	lymphokine-producing macrophage; uncontrolled cell growth; treatment;	
KW	cancer; cytotoxicity; antigen; isoform.	
XX		
OS	Viscum album.	
XX		
FH	Key	Location/Qualifiers
FT	Misc-difference 15	/label= Asp, Glu
FT	Misc-difference 63	/label= Gly, Gln
FT	Misc-difference 66	/label= Ile, Val
FT	Misc-difference 75	/label= Leu, Ala
FT	Misc-difference 107	/label= Asp, none
FT	Misc-difference 113	/label= Asn, Thr
FT	Misc-difference 117	/label= Pro, Thr
FT	Misc-difference 134	/label= Asp, Glu
FT	Misc-difference 141	/label= Ser, Thr
FT	Misc-difference 145	/label= Phe, Tyr
FT	Misc-difference 152	/label= Thr, Ala
FT	Misc-difference 177	/label= Ala, Tyr
FT	Misc-difference 180	/label= Tyr, Asp
FT	Misc-difference 185	/label= Ala, Glu
FT	Misc-difference 191	/label= Val, Met
FT	Misc-difference 219	/label= Ile, Phe
FT	Misc-difference 224	/label= Pro, Ser
FT	Misc-difference 225	/label= Pro, Thr
FT	Misc-difference 232	

FT Misc-difference 236 /label= Thr, Ser
 FT /label= Asp, Ser
 XX DE19804210-A1.
 XX 12-AUG-1999.
 XX
 XX 03-FEB-1998; 98DE-1004210.
 XX
 XX 03-FEB-1998; 98DE-1004210.
 XX
 XX 03-FEB-1998; 98DE-1004210.
 XX
 XX (BIOS-) BIOSYN ARZNEIMITTEL GMBH.
 XX
 XX Morris P, Stiefel T, Voelter W, Welters P;
 XX WPI; 1999-445335/38.
 XX
 XX Preparation of mistletoe lectins in heterologous systems,
 XX particularly for use as anticancer agents and immunostimulants
 XX
 XX Claim 41: Page 39; 78pp; German.

CC This invention describes a novel mistletoe lectin (I) and its fragments
 CC which have antitumour and immunostimulatory activity. The A-chain (MLA)
 CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of
 CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and
 CC lymphokine-producing macrophages, so stimulate immunity. (I) and its
 CC fragments are used to treat uncontrolled cell growth (particularly
 CC cancers) and if they lack cytotoxicity, to increase the strength of the
 CC immune response, particularly to a co-administered antigen
 CC (tumour-associated, bacterial or viral). The method allows production of
 CC mistletoe lectin, and its individual chains, in many different isoforms
 CC and on a large scale, at any time of the year. Recombinant products are
 CC free from toxins present in natural mistletoe extracts. This sequence
 CC represents a consensus sequence of the mistletoe lectin A chain (MLA)
 CC described in the invention.

XX Sequence 255 AA;

alignment_scores:
 Quality: 599.00 Length: 140
 Ratio: 4.538 Gaps: 0
 Percent Similarity: 94.286 Percent Identity: 84.286

alignment_block:

US-09-627-165B-15 x AAY25977

Align seg 1/1 to: AAY25977 from: 1 to: 255

1 TACACAGATCTGAGCGATACCGCGTCATAGGAGCACAGATCCCTGGG 50
 |||:|||||
 116 Tyr***AspLeuGluArgTyrAlaGlyHisArgAspGlnLeuProlu 132
 |||:|||||
 51 TATAGAGGACATCATCCGTCGCGCTTCGTTATCCAGCGCGCA 100
 |||:|||||
 132 Yile***GlnleuilegInserVal***AlaLeuArg***ProGlyIys 149
 |||:|||||
 101 GCACCGGCGCCAAAGCTCGTCCCTTAATCCATCATGATGATCTCC 150
 |||:|||||
 149 ertHrArg***GlnAlaArgSerIleleuIleleuIleGlnMetIleSer 165
 |||:|||||
 151 GAGCGCGGAGATTCATCCCATCTTTGGAGGCGCTCGCCATATCAATTA 200
 |||:|||||
 166 GluAlaAlaArgPheAsnProIleLeuTyrPArg***ArgGln***IleAs 182
 |||:|||||
 201 CAGCGGGAGATCTTCTCCGACATGTACATGCTCGAGCTGGAGACTA 250
 |||:|||||
 182 nserGly***SerPheleuProAsp***TyrMetleuGluLeuGluThrS 199
 |||:|||||
 251 GTTGGGCGCAACATCCAGCAAGTCCAGCATCTACGATGCGGCTTTT 300
 |||:|||||

199 ertPqlyGlnGlnSerThrGlnValGlnHisSerThrAspGlyValPhe 215
 301 AAPAACCCATTTCCGTTGGGTATATCCACCGGTAACTTCGTGACGTGAC 350
 |||:|||||
 216 AsnAsnPro***ArgLeuAlaIle***GlyAsnPheValThrIleu** 232
 351 CAATGTCGCGACGATGATCGCACCTTAGCATCATGTTGTTGTAATGA 400
 |||:|||||
 232 *AsnValArg***ValIleAlaSerLeuAlaIleMetleuPheValCysC 249
 401 GGGACCGACCATCTTCCTCC 420
 |||:|||||
 249 IyGluArgProSerSerSer 255

seq_name: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT: AAB47096

seq_documentation_block:

ID AAB47096 standard; Protein; 120 AA.

AC AAB47096;

DT 16-MAY-2001 (first entry)

DE Korean mistletoe lectin #1.

KW Isoform: A-chain; B-chain; biosynthesis; lectin; Korean mistletoe;

KW KML; tumour; KM-110; KML-C; KMBBP; KML-IIU; KML-IIL;

XX heparin binding protein.

OS Viscum album coloratum.

XX Key Location/Qualifiers

FT Misc-difference 7 /note= "Encoded by NTG"

FT Misc-difference 19 /note= "Encoded by TCN"

FT Misc-difference 64 /note= "Encoded by NTG"

FT Misc-difference 111 /note= "Encoded by NTG"

FT Misc-difference 111 /note= "Encoded by NTG"

FT Misc-difference 111 /note= "Encoded by NTG"

FT Misc-difference 111 /note= "Encoded by NTG"

FT Misc-difference 111 /note= "Encoded by NTG"

FT Misc-difference 111 /note= "Encoded by NTG"

FT Misc-difference 111 /note= "Encoded by NTG"

FT Misc-difference 111 /note= "Encoded by NTG"

FT Misc-difference 111 /note= "Encoded by NTG"

FT Misc-difference 111 /note= "Encoded by NTG"

FT Misc-difference 111 /note= "Encoded by NTG"

FT Misc-difference 111 /note= "Encoded by NTG"

FT Misc-difference 111 /note= "Encoded by NTG"

FT Misc-difference 111 /note= "Encoded by NTG"

FT Misc-difference 111 /note= "Encoded by NTG"

FT Misc-difference 111 /note= "Encoded by NTG"

FT Misc-difference 111 /note= "Encoded by NTG"

FT Misc-difference 111 /note= "Encoded by NTG"

FT Misc-difference 111 /note= "Encoded by NTG"

FT Misc-difference 111 /note= "Encoded by NTG"

FT Misc-difference 111 /note= "Encoded by NTG"

FT Misc-difference 111 /note= "Encoded by NTG"

FT Misc-difference 111 /note= "Encoded by NTG"

FT Misc-difference 111 /note= "Encoded by NTG"

FT Misc-difference 111 /note= "Encoded by NTG"

FT Misc-difference 111 /note= "Encoded by NTG"

FT Misc-difference 111 /note= "Encoded by NTG"

FT Misc-difference 111 /note= "Encoded by NTG"

alignment_scores: Length: 122
 Quality: 408.50
 Ratio: 4.045 Gaps: 1
 Percent Similarity: 82.787 Percent Identity: 69.672

alignment_block:

US-09-627-165b-15 x AAB47096 ..

Align seg 1/1 to: AAB47096 from: 1 to: 120

```

157 GCGAGATTCATCCATCTTTGGAGGCTGCCCATATACATTAAACGGC 206
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
1 AlaArgPheAnProIle***TrpArgLeuAlrArgInIleAsnSerGI 17
207 GGAGCATTTCTTCCCGCATGATCGTACGAGCTGAGAGATGTTGGG 256
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
17 yGInSerSerProProAsnMetIyrmelleuGluLeuGlnThrSerTrpG 34
257 GCCAACAATCCACGCAAGTCGACGAGTCTACGATGGCGTTTAAATAC 306
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
34 LyArgGlnSerThrcInValGlnGlnSerLysAspGlyIlePheAsnThr 50
307 CCATTTCGGTGGTATATCCACCGGTAACTCGAGCGTGAACAATGT 356
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
51 GlnIleArgLeuGlnIleSerAlaGlyAsnPheValThr***SerAsnVa 67
357 TCGCGCGATGATCGCAGCTTAGCATCATGTTGTTGTATGTAAGGACC 406
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
67 lArGAspValIleSerSerLeuAlaIleMetLeuPheGluCysSerGlyA 84
407 GACCATCTTCCTCCGACGTCGCTATGTGGCCGCTGCATACGACCCGTC 456
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
84 rGrProPheSerSerLeuAspHisProSerProLeuLeuAsnGlySerVal 100
457 TTGGAAATATGAGCGCGCTCGACGATGTTACCTGCACGTCTTCCGAAC 506
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
101 ValAsp.....AlaAlaAsnAspValThrCysThr***SerGluPr 114
507 CACCGTCGCGATCGTA 522
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
114 oThValArgIleVal 119

```

seq_name: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT:AAp70097

seq_documentation_block:

ID AAP70097 standard; protein: 332 AA.

```

XX AC AAP70097;
XX XX
XX 09-APR-1991 (first entry)
XX DE Ricin A.
XX XX Ricin A; Met-aminopeptidase.
XX OS Escherichia coli.
XX XX EP219237-A.
XX PD 22-APR-1987.
XX PF 19-SEP-1986; 86EP-0307242.
XX PR 06-MAY-1986; 86US-0860330.
XX PR 20-SEP-1985; 85US-0778414.
XX PA (CERTU ) CERTUS CORP.
XX PI Benbasat A, Bauer KA, Chang S, Chang SY;
XX WPI: 1987-110172/16.
XX DR N-PSDB; AAN70152.
XX

```

PT N-terminal methionine free proteins produ. - by using host
 PT transformed with vector to express a methionine-amino-peptidase
 PS Disclosure; Fig. 4; 20pp; English.

XX Ricin A may be produced in a form which lacks an N-terminal Met
 CC using Met-aminopeptidase from E.coli.
 CC

XX Sequence 332 AA.

alignment_scores: Length: 178
 Quality: 259.50
 Ratio: 2.257 Gaps: 5
 Percent Similarity: 64.607 Percent Identity: 41.011

alignment_block:

US-09-627-165b-15 x AAP70097 ..

Align seg 1/1 to: AAP70097 from: 1 to: 332

```

1 TACACAGATCTGGAGCGATACCGGTCAT...AGGACGAGATCCCTCT 47
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
158 TyAspArgLeuGluGlnLeuAlaGlyAsnLeuArgGlnAsnIleGluLe 174
48 GGGTATAGAGAACTCATTCATCCCTCGCGCTCGCTTAT..... 90
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
174 uGlyAsnGlyProLeuGluGlnAlaIleSerAlaLeuYrTrpYrSerT 191
91 ..CCAGCGCGACGACCGCGGCCCAAGCTCGTCCCTTAATCATCTCAT 138
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
191 hrcIyGlyThrGlnLeuProThrLeuAlaArgSerPheIleIleCysIle 207
139 CAGATATCTCCGAGCGCGGACGATTCATCCATCTTTGGAGGCTCG 188
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
208 GlnMetIleSerGlnAlaAlaArgPheGlnTrpIleGluGluMetAr 224
189 CCATACATTAAACGGGGGAGGATTCCTCTCCGACATGATGCTCG 238
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
224 gThrArgIleArgTrpAsnArgArgSerAlaProAspProSerValIleT 241
239 AGCTGAGACTAGTTGGGCGCAACATCCACGCAAGCTCCAGAGCTTAC 288
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
241 hrlEuGluAsnSerTrpGlyArgLeuSerThrAlaIleGlnGluSerAsn 257
289 GATGCGCTTTTAATACCATTTGGTGGTATATCCACGCGTAATCT 338
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
258 GlnGlyAlaPheAlaSerProIleGlnLeuGlnArgArgAsnGlySerly 274
339 CGTGACGTTGACAAATGTGCGACGATGCCACCTTAGCGATCATGT 388
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
274 sPheSerValTrpAspValSerIleLeuIleProIleIleAlaLeuMetV 291
389 TGTGTATGTAGGACCGACCATCTTCCTCCGACGTCGCGCTTGGCGG 438
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
291 alTYArgCysAlaProProProSerSerGln.....PheSer 303
439 CTGCTCATGACGCGCTTGAATAATAGCGCGCTCGACAGATGTTAC 488
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
304 LeuLeuIleArgProValValProAsnPheAsnAla.....AspVal.. 317
489 CTGCACTGCTTCGAACCCACCGCTGCGCATCGTA 522
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
318 .CysMetAspProGluProIleValArgIleVal 328

```

seq_name: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT:AAp70838

seq_documentation_block:

ID AAP70838 standard; protein: 332 AA.

```

XX AC AAP70838;
XX XX
XX 18-FEB-1991 (first entry)
XX

```


Wed Jul 31 08:46:42 2002

us-09-627-165b-15.ra1

Page 2

401 GGGAGCCGACCATCT 414
:::|||||
231 YGLIATGProser 235

seq_name: /cgn2_6/ptodata/2/1aa/5a.COMB.pep:us-08-776-059-31

seq_documentation_block:

; Sequence 31, Application US/08776059B
; Patent No. 6271368
; GENERAL INFORMATION:
; APPLICANT: LENTZEN, Hans
; APPLICANT: ECK, Jurgen
; APPLICANT: BAUR, Axel
; APPLICANT: ZINKE, Holger
; TITLE OF INVENTION: RECOMBINANT MISTLETOE LECTIN (RML)
; FILE REFERENCE: 674503-2003
; CURRENT APPLICATION NUMBER: US/08/776,059B
; CURRENT FILING DATE: 1999-06-19
; EARLIER APPLICATION NUMBER: PCT/EP96/02273
; EARLIER FILING DATE: 1996-06-25
; EARLIER APPLICATION NUMBER: 95109949.8
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patentn Ver. 2.0
; SEQ ID NO 31
; TYPE: PRT
; ORGANISM: Viscum album
US-08-776-059-31

alignment_scores:

Quality: 619.00 Length: 138
Ratio: 4.654 Gaps: 0
Percent Similarity: 96.377 Percent Identity: 86.232

alignment_block:

US-09-627-165b-15 x US-08-776-059-31 ..

Align seg 1/1 to: US-08-776-059-31 from: 1 to: 253

1 TACACAGATCTGGAGCATACGCGGCATATGAGGACGATCCCTCTGG 50
||| |||||
116 TyrProAspLeuIleuIleGlnSerValThrAlaLeuArgPheProGlyLys 132
51 TATAGAGAACTCATTCATCCGCTCTGCGGCTTCGTATCCAGGGGCA 100
|||||
132 YLLeAspGlnLeuIleGlnSerValThrAlaLeuArgPheProGlyLys 149
101 GCACCCGGGCCAAGCTCGTCCCTTATATCTCATTCAGATGATCTCC 150
|||||
149 erThrArgThrGlnAlaArgSerIleLeuIleGlnMetIleSer 165
151 GAGGCGCGAGATTCATCCATCTTTGGAGGCTGCCAATACATTA 200
|||||
166 GlnAlaAlaArgPheAsnProIleLeuThrPargLaaGlnIleLys 182
201 CACCGGGAGTCACTTCTCCGAGATGATAGCTGAGCTGAGACAA 250
|||||
182 nSerGlyAlaSerPheLeuProAspValTyrMetLeuGlnLeuGlnLys 199
251 GTTGGGGCCAAACATCCACAGAGTCCAGAGTCTACGATGGCGTTT 300
|||||
199 erTrrGlyGlnGlnSerThrGlnValGlnHisSerThrAspGlyValPhe 215
301 AATTAACCATTTGCTGGGTATATCCAGCGTAACTGCTGAGCTTGG 350
|||||
216 AsnAsnProIleArgLeuAlaIleProProGlyAsnPheValThrLeu 232
351 CAATGTCGCGAGCTGATCCGAGCTTATAGCATCATGTTGTTGATG 400
|||||
232 TAsnValArgAspValIleAlaSerLeuAlaIleMetLeuPheValCysG 249

401 GGGAGCCGACCATCT 414
:::|||||
249 YGLIATGProser 253

seq_name: /cgn2_6/ptodata/2/1aa/5a.COMB.pep:us-08-378-761a-77

seq_documentation_block:

; Sequence 77, Application US/08378761A
; Patent No. 5635384
; GENERAL INFORMATION:
; APPLICANT: MALSH, TERENCE A
; APPLICANT: HEY, TIMOTHY D
; APPLICANT: MORGAN, ALICE ER
; TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE
; TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ANDREA T. BORUCKI
; STREET: 9330 ZIONSVILLE ROAD
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: US
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/378,761A
; FILING DATE: 26-JAN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BORUCKI, ANDREA T
; REGISTRATION NUMBER: 33651
; REFERENCE/DOCKET NUMBER: 38272B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317) 337-4846
; INFORMATION FOR SEQ ID NO: 77:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 540 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-378-761a-77

alignment_scores:

Quality: 256.00 Length: 177
Ratio: 2.246 Gaps: 4
Percent Similarity: 64.407 Percent Identity: 39.548

alignment_block:

US-09-627-165b-15 x US-08-378-761a-77 ..

Align seg 1/1 to: US-08-378-761a-77 from: 1 to: 540

1 TACACAGATCTGGAGCATACGCGGCATATGAGGACGATCCCTCTGG 50
||| |||||
123 TyrAspArgLeuGlnGlnIleGlnGlyLeuAlaArgGlnAsnIleGlnLeu 139
51 TATAGAGAACTCATTCATCCGCTCTGCGGCTTCGTAT.....C 91
|||
139 YThrGlyProLeuGlnAspAlaIleSerAlaLeuTyrTyrTyrSerThrC 156
92 CAGGCGCGAGCAACCGCGGCAAGCTGTCCTTATATCCATTCAG 141
|||
156 YSerGlyThrGlnIleProThrLeuAlaArgSerPheMetValCysIleGln 172
142 ATGATCTCCGAGGCGCGAGATTCATTCATCTTTGAGAGGCTCGCA 191

```

|||||
173 MetIleSerGluAlaAlaArgPheGlnTyrIleGluGlyGluMetArgTh 189
192 ATACATTACAGCGGGAGCTATTCTTCCCGACATGTCACATCGCCGAGC 241
189 rArgIleArgTyrAsnArgSerAlaProAspProSerValIleThrL 206
242 TGGAGACTAGTGGGGCCACATCCAGCAGCATGTCAGATTCAGAT 291
206 eugIuAsnSerTrrpGlyArgLeuSerThrAlaIleGlnGluSerAsnGln 222
292 GCGCTTTTATATACCATTTGGTGGGTATATCCACCGTACTTCGT 341
223 GlyAlaPheAlaSerProIleGlnLeuGlnArgAlaGlnGlySerLysph 239
342 GACGTGAGCAATGTTCGGCAGCTGATCGCAGCTTACGATGTTGT 391
239 eAsnValTyrAspValSerIleLeuIleProIleIleAlaLeuMetValT 256
392 TTGTATGAGGACCGACCATCTCTCCGACGTGCGCTATTGGCCGCTG 441
256 yTArgCysAlaProProSerSerGln.....PheSerLeu 268
442 GTCATACGACCCGCTTGGAAATAGCGGGCCGCTCGACGATGTACTGT 491
269 LeuIleArgProValValProAsnPheAsnAla.....AspVal...Cy 282
492 CACTGCTTCCGACCGCTGCGCATCGTA 522
282 sMetAspProGluProIleValArgIleVal 292

seq_name: /cgn2_6/ptodata/2/1aa/5b_COMB.pep:us-08-485-286-77

seq_documentation_block:
: Sequence 77, Application US/08485286
: Patent No. 5646026
: GENERAL INFORMATION:
: APPLICANT: WALSH, TERENCE A
: APPLICANT: HEY, TIMOTHY D
: APPLICANT: MORGAN, ALICE ER
: TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE
: TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF
: TITLE OF INVENTION: USING
: NUMBER OF SEQUENCES: 81
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: ANDREA T. BORUCKI
: STREET: 9330 ZIONSVILLE ROAD
: CITY: INDIANAPOLIS
: STATE: IN
: COUNTRY: US
: ZIP: 46268
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/485,286
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/378761
: FILING DATE: 26-JAN-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: BORUCKI, ANDREA T
: REGISTRATION NUMBER: 33651
: REFERENCE/DOCKET NUMBER: 38272B
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (317) 337-4846
: INFORMATION FOR SEQ ID NO: 77:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 540 amino acids

```

```

: TYPE: amino acid
: STRANDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-485-286-77

alignment_scores:
: Quality: 256.00 Length: 177
: Ratio: 2.246 Gaps: 4
: Percent Similarity: 64.407 Percent Identity: 39.548

alignment_block:
US-09-627-165B-15 x US-08-485-286-77 ..
Align seg 1/1 to: US-08-485-286-77 from: 1 to: 540

1 TACACAGATCTGAGACCGATACCGCGGCTCATATGAGACCAATCCCTGGG 50
|||
123 TyrAspArgLeuGlnGlnLeuGlyGlyLeuArgGluGlnIleGluLeuG 139
51 TATAGAGACTCATTCATCCGCTCGCGCTTCTTAT.....C 91
|
139 yThrGlyProLeuGlnAspAlaIleSerAlaLeuTyrTyrSerThrC 156
92 CAGGGGAGCAGCACCGGCCCAAGCTGCTTCCCTATATATCCATTCAG 141
|||
156 ySgIyThrGlnIleProThrLeuAlaArgSerPheMetValCysIleGln 172
142 ATGATCTCGGAGCGCGGATTCATCCATCTTTGGAGGCTCGCA 191
|||||
173 MetIleSerGluAlaAlaArgPheGlnTyrIleGluGlyGluMetArgTh 189
192 ATACATTACAGCGGGAGCTATTCTTCCGACATGTACATGCTCGAGC 241
189 rArgIleArgTyrAsnArgSerAlaProAspProSerValIleThrL 206
242 TGGAGACTAGTGGGGCCACATCCAGCAGCATGTCAGATTCAGAT 291
206 eugIuAsnSerTrrpGlyArgLeuSerThrAlaIleGlnGluSerAsnGln 222
292 GCGCTTTTATATACCATTTGGTGGGTATATCCACCGTACTTCGT 341
|||||
223 GlyAlaPheAlaSerProIleGlnLeuGlnArgAlaGlnGlySerLysph 239
342 GACGTGAGCAATGTTCGGCAGCTGATCGCAGCTTACGATGTTGT 391
239 eAsnValTyrAspValSerIleLeuIleProIleIleAlaLeuMetValT 256
392 TTGTATGAGGACCGACCATCTCTCCGACGTGCGCTATTGGCCGCTG 441
256 yTArgCysAlaProProSerSerGln.....PheSerLeu 268
442 GTCATACGACCCGCTTGGAAATAGCGGGCCGCTCGACGATGTACTGT 491
269 LeuIleArgProValValProAsnPheAsnAla.....AspVal...Cy 282
492 CACTGCTTCCGACCGCTGCGCATCGTA 522
282 sMetAspProGluProIleValArgIleVal 292

seq_name: /cgn2_6/ptodata/2/1aa/5b_COMB.pep:us-08-356-786-10

seq_documentation_block:
: Sequence 10, Application US/08356786
: Patent No. 5877305
: GENERAL INFORMATION:
: APPLICANT: Huston, James S.
: APPLICANT: Oppermann, Hermann
: APPLICANT: Houston, L. L.
: APPLICANT: Ring, David B.
: TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
: TITLE OF INVENTION: Marker

```



```

? COUNTRY: USA
? ZIP: 60606-6402
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/425,336
? FILING DATE: 18-APR-1995
? CLASSIFICATION: 530
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: 08/064,691
? FILING DATE: 12-MAY-1993
? APPLICATION NUMBER: US 07/901,707
? FILING DATE: 19-JUN-1992
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: US 07/787,567
? FILING DATE: 04-NOV-1991
? ATTORNEY/AGENT INFORMATION:
? NAME: Meyers, Thomas C.
? REGISTRATION NUMBER: P-36,989
? REFERENCE/DOCKET NUMBER: 31394
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 312/474-6300
? TELEFAX: 312/474-0448
? TELEX: 25-3856
? INFORMATION FOR SEQ ID NO: 1:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 267 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-425-336-1

```

```

alignment_scores:
? Quality: 217.00 Length: 143
? Ratio: 2.309 Gaps: 2
? Percent Similarity: 65.734 Percent Identity: 39.860

```

alignment_block:

US-09-627-165b-15 x US-08-425-336-1 ..

Align seg 1/1 to: US-08-425-336-1 from: 1 to: 267

```

1 TACACAGATCTGGAGCAATACCGCGCAT...AGGACGACATCCCTCT 47
  ||| |||||::: |||||::: |||::: ||| ||
123 TTTsptrglenulnleuAlaIleSerAlaIleuTYTYTYSerT 139
  ||| |||::: |||::: ||||| |||
48 GGGTATAGAGGAATCATTCATTCGCTCGCGCTTCTTAT..... 90
  ||| |||::: |||::: ||||| |||
139 uGIYasnGIYProLeuGIuGIuAlaIleSerAlaIleuTYTYTYSerT 156
  ||| |||::: |||::: ||||| |||
91 ..CCAGCGGACGACACCGCGGCAAGCTGTTCCCTTATATCCCAT 138
  |||::: |||::: |||||::: ||||| |||
156 hrGIYgIYhrGlnleuProthrleuAlaIleSerPheIleIleGly 172
  |||::: |||::: |||||::: ||||| |||
139 CAGATGATCTCCGAGCGCGGAGATTCATCCATCTTTGGAGGCTCG 188
  ||||| ||||| ||||| ||||| |||
173 GlnleIleSerGIuAlaIleAlaIleArgPheGIuIleIleGIuIle 189
  ||||| ||||| ||||| ||||| |||
189 CCATATCATTTACACCGGAGGAGTCAATTCCTCCGACATGTACATCT 238
  |||::: |||::: ||||| ||||| |||
189 gThrArgIleArgTYAsnArgArgSerAlaProAspProSerValIle 206
  ||||| ||||| ||||| ||||| |||
239 AGGTGAGACTAGTTGGGCGCAACATCCAGCAAGCTCCAGAGTCTAG 288
  ||||| ||||| ||||| ||||| |||
206 hrleuGIuAsnSerITrgIYArgleuSerThrAlaIleGIuIleGly 222
  ||||| ||||| ||||| ||||| |||
289 GATGGCGTTTAAATACCATTCCTGGTGGTATATCCACCGTACTT 338
  ::|||::|||::|||::|||::|||::|||::|||::|||::|||
223 GlnleYAlaIlePheAlaSerProIleGIuIleuGIuIleArgIYasn 239

```

```

339 CGTGACGTGGACATGTTCGCGACGTGATCCGACGATAGCATCATGT 388
  ::|||::|||::|||::|||::|||::|||::|||::|||
239 sPheSerValIYTrspAlaSerIleleuIlePheIleIleAlaIleu 256
  ::|||::|||::|||::|||::|||::|||::|||
389 TGTGTGTATGTAGGACGACATCTTCC 417
  ::|||::|||::|||::|||::|||::|||
256 alTYrtrGcYsAlaProProIleSer 265

```

seq_name: /cgn2_6/prodata/2/1aa/5A_COMB.pep:US-08-488-113B-1

seq_documentation_block:

```

? Sequence 1, Application US/08488113B
? Patent No 5744580
? GENERAL INFORMATION:
? APPLICANT: Better, Marc D.
? APPLICANT: Carroll, Stephen F.
? APPLICANT: Studinka, Gary M.
? TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
? NUMBER OF SEQUENCES: 169
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: McAndrews, Held & Malloy, Ltd.
? STREET: 500 West Madison Street, 34th floor
? CITY: Chicago
? STATE: Illinois
? COUNTRY: USA

```

```

? ZIP: 60661
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/488,113B
? FILING DATE: 07-JUN-1995
? CLASSIFICATION: 530

```

```

? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: US 08/425,336
? FILING DATE: 18-APR-1995
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: US 08/064,691
? FILING DATE: 12-MAY-1993
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: US 07/988,430
? FILING DATE: 09-DEC-1992
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: US 07/901,707
? FILING DATE: 19-JUN-1992
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: US 07/787,567
? FILING DATE: 04-NOV-1991
? ATTORNEY/AGENT INFORMATION:
? NAME: McNicholas, Janet M.
? REGISTRATION NUMBER: 32,918
? REFERENCE/DOCKET NUMBER: 11022US07/200-70.p3.C2A
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 312/707-8889
? TELEFAX: 312/707-9155
? TELEX: 650 388-1248
? INFORMATION FOR SEQ ID NO: 1:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 267 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-488-113B-1

```

```

alignment_scores:
? Quality: 217.00 Length: 143
? Ratio: 2.309 Gaps: 2
? Percent Similarity: 65.734 Percent Identity: 39.860

```


alignment block:

US-09-627-165B-15 x US-08-488-113B-1 ..

Align seg 1/1 to: US-08-488-113B-1 from: 1 to: 267

```

1 TACACAGATCTGAGCGATACGCCGCTCAT...AGGAGCCAGATCCCTCT 47
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
123 TyrAspArgLeuGlnLeuAlaGlnLeuAlaArgGlnAsnIleGluLe 139
48 GCGATAGAGAACTCATTCATCCGCTCGCGCTTCCTAT..... 90
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
139 uGlyAsnGlyProLeuGlnLeuAlaIleSerAlaLeuTyrTyrSerT 156
91 ..CCAGCGCGAGCAGCCGCCGCAATCTGCTTCCTTATATCCCAT 138
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
156 hrcGlyGlyThrGlnLeuProThrLeuAlaArgSerPheIleIleCysIle 172
139 CAGATGATCTCGAGCGCGCGAGATTCATCCATCTTTGGAGGCTCG 188
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
173 GlnMetIleSerGlnAlaAlaArgPheGlnTyrIleGlnGlyGluMetAr 189
189 CCAATACATTAAACAGCGGGGAGTCATTTCTCCGACATGATGATGCTG 238
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
189 gThrArgIleArgTyrAsnArgSerAlaProAspProSerValIleT 206
239 AGCTGAGACTAGTGGGGCCAAACAATCCAGCAAGTCACAGCTACG 288
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
206 hrcLeuGlnAsnSerTpgIlyArgLeuSerThrAlaIleGlnGluSerAsn 222
289 GATGCGCTTTTAAATACCATTTTCGGTGGTATATCCACCGCTACT 338
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
223 GlnGlyAlaPheAlaSerProIleGlnLeuGlnArgArgAsnGlySerLy 239
339 CGTGACGCTAGCAATGTCGCGAGGTGATCCGACGTTAGCATGATG 388
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
239 sPheSerValTyrAspAlaSerIleLeuIleProIleIleAlaLeuMetV 256
389 TGTGTATGTAGGAGCCGACCATCTTC 417
256 aTyrArgCysAlaProProSerSer 265

seq_name: /cgn2_6/ptodata/2/1aa/5a_COMB.pep:US-08-477-484B-1

seq_documentation_block:
: Sequence 1, Application US/08477484B
: Patent No. 5756699
: GENERAL INFORMATION:
: APPLICANT: Better, Marc D.
: APPLICANT: Carroll, Stephen F.
: APPLICANT: Studnika, Gary M.
: TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
: TITLE OF INVENTION: Proteins
: NUMBER OF SEQUENCES: 169
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: McAndrews, Held & Malloy, Ltd.
: STREET: 500 West Madison Street, 34th floor
: CITY: Chicago
: STATE: Illinois
: COUNTRY: USA
: ZIP: 60661
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DO5
: SOFTWARE: Patentn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA: US/08/477,484B
: APPLICATION NUMBER: US/08/477,484B
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/425,336
: FILING DATE: 18-APR-1995

```

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/064,691

FILING DATE: 12-MAY-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/938,430

FILING DATE: 09-DEC-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/901,707

FILING DATE: 19-JUN-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/787,567

FILING DATE: 04-NOV-1991

ATTORNEY/AGENT INFORMATION:

NAME: McNicholas, Janet M.

REGISTRATION NUMBER: 32,918

REFERENCE/DOCKET NUMBER: 11022US07/200-70.P3.C2A

TELEPHONE: 312/707-9155

TELEFAX: 312/707-8889

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 267 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-477-484B-1

alignment_scores:

Quality: 217.00

Ratio: 2.309

Percent Similarity: 65.734

Length: 143

Gaps: 2

Percent Identity: 39.860

alignment block:

US-09-627-165B-15 x US-08-477-484B-1 ..

Align seg 1/1 to: US-08-477-484B-1 from: 1 to: 267

```

1 TACACAGATCTGAGCGATACGCCGCTCAT...AGGAGCCAGATCCCTCT 47
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
123 TyrAspArgLeuGlnLeuAlaGlnLeuAlaArgGlnAsnIleGluLe 139
48 GCGATAGAGAACTCATTCATCCGCTCGCGCTTCCTAT..... 90
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
139 uGlyAsnGlyProLeuGlnLeuAlaIleSerAlaLeuTyrTyrSerT 156
91 ..CCAGCGCGAGCAGCCGCCGCAATCTGCTTCCTTATATCCCAT 138
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
156 hrcGlyGlyThrGlnLeuProThrLeuAlaArgSerPheIleIleCysIle 172
139 CAGATGATCTCGAGCGCGCGAGATTCATCCATCTTTGGAGGCTCG 188
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
173 GlnMetIleSerGlnAlaAlaArgPheGlnTyrIleGlnGlyGluMetAr 189
189 CCAATACATTAAACAGCGGGGAGTCATTTCTCCGACATGATGATGCTG 238
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
189 gThrArgIleArgTyrAsnArgSerAlaProAspProSerValIleT 206
239 AGCTGAGACTAGTGGGGCCAAACAATCCAGCAAGTCACAGCTACG 288
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
206 hrcLeuGlnAsnSerTpgIlyArgLeuSerThrAlaIleGlnGluSerAsn 222
289 GATGCGCTTTTAAATACCATTTTCGGTGGTATATCCACCGCTACT 338
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
223 GlnGlyAlaPheAlaSerProIleGlnLeuGlnArgArgAsnGlySerLy 239
339 CGTGACGCTTAGCAATGTCGCGAGGTGATCCGACGTTAGCATGATG 388
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
239 sPheSerValTyrAspAlaSerIleLeuIleProIleIleAlaLeuMetV 256
389 TGTGTATGTAGGAGCCGACCATCTTC 417
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


REFERENCE/DOCKET NUMBER: DJB/9901/215431/TGM
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3000
TELEFAX: 202-822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 61:
SEQUENCE CHARACTERISTICS:
LENGTH: 267 amino acids
TYPE: amino acid
STRANDEDNESS: Single
TOPOLOGY: Linear
US-08-338-793D-61

alignment_scores: Length: 143
 Quality: 217.00 Gaps: 2
 Ratio: 2.309
Percent Similarity: 65.734 Percent Identity: 39.860

alignment_block:

US-09-627-165b-15 x US-08-338-793D-61 ..

Align seg 1/1 to: US-08-338-793D-61 from: 1 to: 267

```
1 TACACAGATCTGGAGCGATACGCGGTCAAT..AGGACAGATCCCTCT 47
   ||| |||||: |||||: |||||: |||||: |||||: |||||: ||
123 TyrAspArgLeuGluInLeuAlaIleSerAlaLeuTyrTyrSerT 139
   ||| |||||: |||||: |||||: |||||: |||||: |||||: ||
48 GGGTATAGAGGAATCATTCATCCGCTCGGCGCTCGTAT..... 90
   ||| ||| |||: |||||: |||||: |||||: |||||: |||||: ||
139 uGlyAsnGlyProLeuGluGluAlaIleSerAlaLeuTyrTyrSerT 156
   ||| |||: |||: |||: |||: |||: |||: |||: |||
91 ..CCAGCGCGACACCGCGGCCCAAGCTGTCCTTATATCCCTCAT 138
   ||| |||: |||: |||: |||: |||: |||: |||: |||
156 hrGlyGlyThrGlnLeuProThrLeuAlaArgSerPheIleIleCysIle 172
   ||| |||: |||: |||: |||: |||: |||: |||: |||
139 CAGATGATCTCCGAGCGCGGAGATTCATCCATCTTTGAGAGGCTCG 188
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
173 GlnMetIleSerGluAlaAlaArgPheGlnTyrIleGluGlyGluMetAr 189
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
189 CCAATACATTAACAGCGCGGAGTCATTTCTCCGACATGTACATGCTCG 238
   ||| |||: |||: |||: |||: |||: |||: |||: |||
189 gThrArgIleArgTyrAsnArgArgSerAlaProAspProSerValIleT 206
   ||| |||: |||: |||: |||: |||: |||: |||: |||
239 ACCTGGAGACTAGTTGGGCGCAACAATCCAGCAAGTCCAGAGTCTACG 288
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
206 hrLeuGluAsnSerTyrPcIlyArgLeuSerThrAlaIleGlnGluSerAsn 222
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
289 GATGGCGTTTAAATACCAATTCGTTGGGTATATCCACCGGTAACTT 338
   ||| |||: |||: |||: |||: |||: |||: |||: |||
223 GlnGlyAlaPheAlaSerProIleGlnLeuGlnArgArgAsnGlySerLy 239
   ||| |||: |||: |||: |||: |||: |||: |||: |||
339 CGTGACGTTGACCAATGTTCCGAGCGGTGATGCCAGTTCAGATCATGT 388
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
239 sPheSerValTyrAspValSerIleLeuIleProIleIleAlaLeuMetV 256
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
389 TGTTGTATGTAGGACGACGACCATCTCTCC 417
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
256 aIlyArgCysAlaProProProSerSer 265
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```


OW of: US-09-627-165b-15 to: PIR_71: * out_format : pfs
Date: Jul 30, 2002 3:55 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODE=frame+zip,model -DEV=zip
-O=/cgn2.1/USPTO.spool/US09627165/rnat-30072002_151139_3235/alpha_query.fasta_1.1012
-OB=PIR_71 -QFMT=fastan -SUFFIX=rrp -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPEL=0.000 -LOOPEXT=0.000 -GAPOP=4.500
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -DELAP=6.000
-FEAPEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500 -TRANS=human40.cdl
-DEEXT=7.000 -START=1 -MATRIX=biosum62 -TRANS=human40.cdl
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pis -NORM=ext -HEAPSIZE=500
-MINLEN=0 -MAXLEN=200000000 -USER=US09627165_@CGN1_1.172
-NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMECUT=120 -WARN_TIMECUT=30
-NO_XLPHY -WAIT -THREADS=1

Search information block:

Query: US-09-627-165b-15
Query length: 522
Database: PIR_71: *
Database sequences: 283138
Database length: 96089334
Search time (sec): 72.560000

Sequence	Strd	Orig	Zscore	Escore	Len	Documentation
PIR2:PD0018	+	621.00	1152.96	1.6e-56	254	mistletoe lectin I A chain - Vi
PIR1:TTZLSA	+	283.00	510.91	4.3e-21	528	abrin d precursor - Indian li
PIR1:R1CSD	+	259.50	465.94	1.3e-18	576	ricin d precursor - castor bea
PIR2:S32331	+	256.00	460.17	2.9e-18	528	abrin d precursor - Indian li
PIR2:S16022	+	256.00	459.59	2.9e-18	562	abrin c precursor - Indian li
PIR1:RLCSAG	+	256.00	459.56	2.9e-18	564	agglutinin precursor - castor
PIR2:S32430	+	246.50	442.34	2.9e-17	527	abrin b precursor - Indian li
PIR2:S62627	+	237.00	423.75	2.9e-16	570	agglutinin I precursor - Europ
PIR2:C39761	+	220.00	393.47	1.5e-14	251	abrin (clone 7.2) precursor - I
PIR2:R39761	+	180.50	333.29	1.7e-10	106	abrin (clone 3.7) precursor - I
PIR1:RLUTYM	+	163.50	279.75	6.0e-08	289	RNA N-glycosidase (EC 3.2.2.2
PIR2:UC5606	+	156.50	279.77	6.6e-08	261	antiviral protein - Virginian B
PIR2:DE0401	+	156.00	277.71	7.6e-08	294	RNA N-glycosidase (EC 3.2.2.2
PIR2:S28421	+	154.50	276.52	1.1e-07	286	karsurin-B - Trichosanthes kid
PIR2:UC5032	+	154.50	275.15	1.1e-07	247	RNA N-glycosidase (EC 3.2.2.2
PIR1:RLPUGG	+	152.50	272.76	1.7e-07	247	karsurin - Mongolian snake-ven
PIR2:S17787	+	144.50	255.52	1.2e-06	313	RNA N-glycosidase (EC 3.2.2.2
PIR2:S55560	+	141.00	249.78	2.8e-06	286	RNA N-glycosidase (EC 3.2.2.2
PIR2:A59817	+	140.00	248.17	3.6e-06	278	RNA N-glycosidase (EC 3.2.2.2
PIR2:JC4840	+	138.00	245.59	5.6e-06	245	RNA N-glycosidase (EC 3.2.2.2
PIR2:JC4840	+	137.00	241.33	7.4e-06	316	RNA N-glycosidase (EC 3.2.2.2
PIR2:JC4235	+	136.00	240.39	9.4e-06	286	RNA N-glycosidase (EC 3.2.2.2
PIR2:S24294	+	125.00	220.86	0.0001	253	RNA N-glycosidase (EC 3.2.2.2
PIR2:S28539	+	124.00	217.93	0.0002	283	RNA N-glycosidase (EC 3.2.2.2
PIR2:S05205	+	124.00	217.64	0.0002	292	RNA N-glycosidase (EC 3.2.2.2
PIR1:RLDHG2	+	123.50	217.16	0.0002	278	beta-luffin - smooth loofah
PIR2:S23519	+	123.00	217.10	0.0006	253	RNA N-glycosidase (EC 3.2.2.2
PIR2:S28542	+	119.00	209.58	0.0006	253	RNA N-glycosidase (EC 3.2.2.2
PIR2:S29331	+	116.50	204.21	0.0010	272	beta-luffin - beet
PIR2:JC4811	+	114.00	200.30	0.0018	250	luffin-b - smooth loofah
PIR2:JUN0108	+	112.00	195.06	0.0031	293	RNA N-glycosidase (EC 3.2.2.2
PIR2:S17519	+	103.50	181.10	0.0228	236	RNA N-glycosidase (EC 3.2.2.2
PIR2:S17932	+	99.00	172.00	0.0683	253	RNA N-glycosidase (EC 3.2.2.2
PIR2:A58923	+	93.50	159.52	0.2693	318	Shiga-like toxin II chain A pre
PIR2:S01032	+	93.50	159.50	0.2693	319	hypothetical protein - Escheid
PIR2:S21940	+	93.50	159.50	0.2693	319	shiga-like toxin II chain A - B
PIR2:I54695	+	93.50	159.50	0.2693	319	variant shiga-like toxin A subu
PIR2:I76713	+	93.50	159.50	0.2693	319	variant shiga-like toxin A subu

PIR2:E90779 + 93.50 159.50 0.2693 319 | Shiga toxin 2 subunit A [imp
PIR2:685640 + 93.50 159.50 0.2693 319 | Shiga toxin 2 subunit A [imp
PIR2:S58343 + 92.50 157.62 0.3427 319 | Shiga-like toxin II chain A
PIR2:I60446 + 91.50 155.74 0.4361 319 | Shiga-like cytotoxin subunit
PIR2:T12573 + 91.00 155.72 0.4823 289 | RNA N-glycosidase (EC 3.2.2.2

seq_name: PIR2:PD0018

seq_documentation_block:

mistletoe lectin I A chain - Viscum album (fragment)
C:Species: Viscum album
C:Date: 03-Jul-1998 #sequence_revision 10-Jul-1998 #text_change 26-Aug-1999
C:Accession: PD0018
R:Eschenburg, S.; Krauspenhaar, R.; Mikhailov, A.; Stoeva, S.; Betzel, C.; Voelker, W
Biochem. Biophys. Res. Commun. 247, 367-372, 1998
A:Title: Primary structure and molecular modelling of mistletoe lectin I from Viscum a
A:Accession number: PD0018; MUID:98308123
A:Reference: 1-254 <ESC>
A:Molecule type: protein
A:Residues: 1-254 <ESC>
C:Superfamily: ricin; RNA N-glycosidase homology <RNG>
F:7-246/Domain: RNA N-glycosidase homology <RNG>

alignment_scores:

Quality: 621.00 Length: 140
Ratio: 4.634 Gaps: 0
Percent Similarity: 95.714 Percent Identity: 85.714

alignment_block:

US-09-627-165b-15 x PD0018 ..

Align seg 1/1 to: PD0018 from: 1 to: 254

1 TACACAGATCTGAGAGATACGCGGTCATAGGACGACGATTCCTGGG 50
|||||
115 TYPProAspLeuGlnArgTyrAlaGlyHisArgAspGlnIleProLeuG 131
51 TATAGAGAACTCATTCATCGTCGCGGCGCTTGATATCAGGCGCA 100
|||||
131 YLEAspGlnIleIleGlnSerValThrAlaLeuArgProGlyGlyS 148
101 GCACCGCGCGCCAGCTGCTGCTTATATATCCATTCATGATATGCC 150
|||||
148 ERTThrArgThrGlnAlaArgSerIleLeuIleLeuIleIleSer 164
151 GAGCGCGCGATTCATTCATTCCTTTGGAGGCGTCGCATATATTA 200
|||||
165 GlnAlaAlaArgPheAsnProIleLeuTyrArgGlnIleLeuS 181
201 CAGCGGGAGATCATTTCTCCGACATATGATATGATGAGCTGAGACTA 250
|||||
181 NSeGlyAlaSerPheLeuProAspValTyrMetLeuGlnLeuGlnThrS 198
251 GTTGGCGCCAAATTCACAGCAGATGTCAGATGTCAGATGTCGTTT 300
|||||
198 ERTPrGlyGlnGlnSerThrGlnValGlnHisSerThrAspGlyValPhe 214
301 AATACCATTCGTTGGTATATCCACCGATTCGTAATTCGTAATTCG 350
|||||
215 AsnAsnProIleArgLeuAlaIleProProGlyAsnPheValThrLeuTh 231
351 CAATGTCGCGAGTATGTCGCGATTCAGCATATATTCGTTTGTATGA 400
|||||
231 RAsnValArgAspValIleAlaSerLeuAlaIleMetLeuPheValGlyS 248
401 GGGAGCGACATTCCTCTC 420
|||||
248 LyeGlnArgProSerSerSer 254
seq_name: PIR1:TTZLSA
seq_documentation_block:

C: Superfamily ricin; rRNA N-glycosidase homology
C: Keywords: duplication; glycoprotein; glycosidase; hydrolase; lectin; pyroglutamic
F:1-212/Product: abrin-a chain A #status experimental <ACH>
F:7-246/Domain: rRNA N-glycosidase homology <RNC>
F:261-528/Product: abrin-a chain B #status experimental <RCH>
F:283-325,326-366,369-407,414-449,453-492,495-528/Region: 40-residue repeats
F:1/Modified site: pyroglutamate carboxylic acid (Glu) #status experimental
F:74,113,195,196/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted
F:164,167/Active site: Glu, Arg #status predicted
F:247-269,286-305,329-346,417-430,456-473/Disulfide bonds: #status predicted
F:288,312/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted
F:361,401/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted
F:500,521/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted

Alignment_scores:
Quality: 283.00 Length: 185
Ratio: 2.320 Gaps: 6
Percent Similarity: 65.946 Percent Identity: 38.919

Alignment_block:
US-09-627-165B-15 x TZLSA ..

Align seg 1/1 to: TZLSA from: 1 to: 528

```
1 TACACAGATCTGGAGCCGATACGCCGTCAT...AGGACACAGATCCCTC 47
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
113 Tyrolspsrleugluatrgtprlalnlglnseratrglnlgnllepole 129
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
48 GGTATATAGAGAACTCATTCATCCGTCGCCGCTTCGTTATTCAGAGCG 97
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
129 uglyleuglnalaleuthrhnselgylileserphieatrgclgclya 146
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
98 GCAGCACCAGCCGCCCAAGCTGTCCTCTTAATTCCTCATCAGATGATC 147
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
146 snspansngluilgylsalatrgthreullelvalileleglmetval 162
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
148 TCCGAGGCCCCGAGATTCATTCCTCTTTGGAGGGCTCCCAATACAT 197
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
163 Alagluhlaialarphcatgyltlileseranlatvalatvalserll 179
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
198 TAACACCGGGGAGCTCATTTCTTCCGACATGTACATGCTCGAGCTGAGA 247
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
179 eglntprgltgltthralaPhglnProspalaaImetlileserleugl 196
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
248 CTAATGGGGGCCAACAATCCAGCAAGTCCAGCAAGTCCAGTACGATGCG 297
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
196 snasntprspasnleuseratrgclgylalnglnluseratglnlspr 212
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
298 TTTAATAACCATTTCCGTTGGGTATATCCAGCCGTAACCTTCGAGCGT 347
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
213 PhieProasnGln.....ValThrLe 219
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
348 GAGCAATCTTCCGAC.....G 364
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
219 uthrlnlileatrgasngluprovalilevalaspsrleuserhlsprot 236
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
365 TGAATGCCAGCTTAGCGATCATGTGTTGTATGTAGGACCGACCATCT 414
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
236 hvalaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 251
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
415 TCTCTCCGACGTGCGCTATTTGCCGCTGTACAGCCGCTCTTGAAAA 464
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
252 Alaasnglnser.....Proleuileuletergserllevalgluly 265
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
465 TAGCGGGCGCGCTCGACATGTATCTGCACACTGCTTCCGAACCCAGCT 514
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
265 sSer.....LysilecysSerSeratrgtlygluprothvala 278
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
515 GCATC 519
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
278 rgile 279
```


BioSci. Biotechnol. Biochem. 57, 166-169, 1993
 A:Title: The complete amino acid sequences of the B-chains of Abirin-a and Abirin-b, toxic
 A:Reference number: JCI398; MUID:93169023
 A:Accession: JCI399
 A:Molecule type: protein
 A:Residues: 260-281, 'D', 283-290, 'N', 292-349, 'PQ', 352-377, 'N', 379-425, 'W', 427, 'D', 429-430
 A:Experimental source: seed
 C:Keywords: ricin; rRNA N-glycosidase homology
 F:1-350/Product: abirin-b chain A #status predicted
 F:7-243/Domain: rRNA N-glycosidase homology <RNG>
 F:260-527/Product: abirin-b chain B #status experimental <BGH>
 F:262-324, 325-365, 368-406, 413-448, 452-491, 494-527/Region: 40-residue repeats
 F:74, 113, 194, 195/Binding site: pyroglutamate carboxylic acid (Glu) #status predicted
 F:110, 360, 400/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted
 F:163, 166/Active site: Glu, Arg #status predicted
 F:246-268, 285-304, 328-345, 416-429, 455-472/Disulfide bonds: #status predicted
 F:287, 311/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted
 F:499, 520/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted

alignment_scores:
 Quality: 246.50 Length: 175
 Ratio: 2.071 Gaps: 6
 Percent Similarity: 68.000 Percent Identity: 38.857

alignment_block:
 US-09-627-165b-15 x S32430 ..

Align seg 1/1 to: S32430 from: 1 to: 527

```

1 TACACAGATCTGGAGCGATACCGCGTCAT...AGGACACAGATCCCTCT 47
  ||| ||||| ||||| ||| :: ||||| ||||| |||||
113 TyrIleAspIleuGlnArgIleuGlnThrArgIleuGlnIleProle 129
48 GGGTATAGAGAACTCATTCATCCGCTCGCGCTTCGTTATCCAGCG 97
  ||||| ::||| ::||| ::||| ::||| ::|||
129 uGlyIleuGlnAlaIleuArgIleuGlnIleSerPheIleuGln...SerGlyT 145
98 GCACGACCCGGGCCCAAGCTCGTCCCTTATATCCATTCATTCATATC 147
  :: :: ||||| ::||| ::||| ::||| ::|||
145 hrAspAspIleuGlnIleAlaArgPheIleuIleValIleIleGlnMetAla 161
148 TCCGAGCGCGGAGATTCATCCATCCATCTTTTGGAGGCGTCCCAATACAT 197
  ||||| ||||| ||||| ::||| ::||| ::|||
162 SerGlnAlaIleuArgTyrArgPheIleSerTyrArgValIleGlyValSerIle 178
198 TAACAGCGGAGATCATTTCTCCGACATGTATACATCTGAGCTGAGAGA 247
  ||||| ::||| ||||| ||||| ||||| |||||
178 earGThrAsnThrAlaPheGlnProAspAlaIleMetIleSerIleuGln 195
248 CTAGTTGGGGCCACAAATCCACGAACTCAGAGTCTACGATGCGCTT 297
  ::||| ::||| ||||| ||||| ||||| ::|||
195 snAsnTTPAspAsnIleuSerGlyIleValIleGlnIleSerValIleGlnAspThr 211
298 TTTAATACCAATTCGTTG...GGATATCCACCGGTAATCTCGTGAC 344
  ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
212 PheProAsnAlaValIleuArgSerValAsnAsnGlnProValIleVal 228
345 GTTGAGCAATGTTCCGAGCTGATCCAGCTTAGCATGATGTTGTTG 394
  ::||| ::||| ||||| ||||| ||||| |||||
228 IAspSerIleuThrHisGlnSerValAlaValIleuAlaIleuMetIlePheV 245
395 TATGTAGGACGACCATCTTCCTCCGAGTCCGCTATTTGGCGCTGTG 444
  ||||| ::||| ||||| ::||| ||||| ::|||
245 alGys...AsnProProAsnAlaAsnGlnSer.....ProIleu 257
445 ATACGACCCGCTTTGGAAATAGCGCGCGCTGACGATGTACTGTGAC 494
  ||||| ::||| ||||| ::||| ||||| ::|||
258 IleArgSerIleValIleuIleuSer.....LysIleGlySerSer 270
495 TGCTTCGAACCCACGCTGCGATC 519
```

seq_name: p1r2:S62627
 270 rArgTyrGlnProThrValArgIle 278
 ||||| ||||| ||||| ||||| ||||| ||||| |||||

seq_documentation_block:
 agglutinin I precursor - European elder
 C:Species: Sambucus nigra (European elder)
 C:Date: 24-Aug-1996 #sequence
 C:Accession: S62627; S62619
 R:van Damme, E.J.M.; Barre, A.; Rouge, P.; van Leuven, F.; Peumans, W.J.
 Eur J. Biochem. 235, 128-137, 1996
 A:Title: The NeuAc(alpha-2,6)-Gal/galNAc-binding lectin from elderberry (Sambucus nigra)
 A:Reference number: S62619; MUID:96202926
 A:Accession: S62627
 A:Status: preliminary
 A:Molecule type: rRNA
 A:Residues: 1-570 <YAN>
 A:Cross-references: EMBL:U27122; NID:q1141772; PIDN:AMC49158.1; PID:q1141773
 A:Accession: S62619
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 29-39;309-319 <VA2>
 C:Superfamily: ricin; rRNA N-glycosidase homology
 F:37-283/Domain: rRNA N-glycosidase homology <RNG>

alignment_scores:
 Quality: 237.00 Length: 186
 Ratio: 1.911 Gaps: 7
 Percent Similarity: 66.667 Percent Identity: 34.946

alignment_block:
 US-09-627-165b-15 x S62627 ..

Align seg 1/1 to: S62627 from: 1 to: 570

```

1 TACACAGATCTGGAGCGATACCGCGTCAT...AGGACACAGATCCCTCT 47
  ||||| ::||| ::||| ::||| ::||| ::|||
145 TyrThrIleuGlnArgIleuGlnValIlePheGlyArgValIleThrIleProle 161
48 GGGTATAGAGAACTCATTCATCCGCTCGCGCTTCGTTATCCAGCG 97
  ||||| ::||| ||||| ||||| ||||| |||||
161 uGlyProIleuSerIleuAspGlnAlaIleSerIleuArgThrIleThrIle 178
89 ATCCAGCGGCGACCGCGGCCCAAGCTCGTCCCTTATATCCATTCAT 138
  ::||| ::||| ||||| ||||| ||||| |||||
178 euThrAlaGlyAspThrIleuProIleuAlaArgIleuIleuValIle 194
139 CAGATGATCTCCGAGCGGAGATTCATCCATCCATCTTTTGGAGGCGTGC 188
  ||||| ::||| ||||| ||||| ||||| |||||
195 GlnMetValSerGlnAlaIleArgPheArgTyrIleGlnIleuArgIleu 211
189 CCATACATATACAGCGGAGATTCATTTCTCCGACATGTACATGCTCG 238
  ||||| ::||| ||||| ||||| ||||| |||||
211 gThrSerIleThrAspAlaSerGlnIleuThrProAspIleuMetIleu 228
239 AGCTGAGACTAGTTGGGGCCACAAATCCACGACAGTCCAGAGCT... 285
  ::||| ::||| ||||| ||||| ||||| |||||
228 erMetGlnAsnAsnIlePheSerSerMetSerSerGlnIleuIleGlnAlaIle 244
286 ACGGATGGCGTTTATATACCAATTCGTTGGGTATATCCACCGGTAA 335
  ||||| ::||| ||||| ::||| ||||| ::|||
245 ProGlyIleIlePheAlaGlyValAlaGlnIleuArgAspGluArgAsnAs 261
336 CTTCGTGAGGTTGAGCAATGTTCGGGAGCTG...ATGCCAGCTTAG 379
  ::||| ::||| ||||| ::||| ||||| ::|||
261 nSerIleGlnValIleThrAsnPheArgIleuPheGlnIleuThrTyrIle 278
380 CGATCATGTGTTTGTATGTAGGACGACCATCTTCCTCC... 420
  ||||| ::||| ||||| ::||| ||||| ::|||
278 IeValIleuIleuTyrGlyAsnAlaProValIleThrSerSerSerTyrSerAsn 294
```

421GACGTGCGCTATGCGCGCTGTCATACGACCCGCTTTGGA 461
|||||
295 AsnAlaIeIspAlaIeIn.....IleIleIysMeIeProAlaIeAheI 308
462 AAATAGCGCGCGCGTGCAGCATTTACCTGCTCCGGAACCAACCG 511
: |||
308 g.....GlyGlyGlyIuTyrGluIuIysValCysSerValIaIuIValThrA 323
512 TCGCGATC 519
|||||
323 rgaIgtIle 325

seq_name: p1r2:C39761

seq_documentation_block:
abrin (clone 7.2) precursor - Indian licorice (fragment)
N:contains: rRNA N-glycosidase (EC 3.2.2.22)
C:Species: Abrus precatorius (Indian licorice)
C:Date: 21-Feb-1992 #sequence_revision 31-Dec-1993 #text_change 20-Aug-1999
C:Accession: C39761; S14471
R:Evensen, G.; Mathiesen, A.; Sundan, A.
J. Biol. Chem. 266: 6848-6852, 1991
A:Title: Direct molecular cloning and expression of two distinct abrin A-chains.
A:Reference number: A39761; MUID:91201329
A:Accession: C39761
A:Molecule type: DNA
A:Residues: 1-251 <EVE>
R:Evensen, G.; Mathiesen, A.; Sundan, A.
submitted to the EMBL Data Library, October 1990
A:Description: Direct molecular cloning of two distinct abrin A-chains.
A:Reference number: S14471
A:Accession: S14471
A:Molecule type: DNA
A:Residues: 'M', 1-251 <EVE>
A:Cross-references: EMBL:554872; NID:g16088; PIDN:CAA38654.1; PID:g16089
C:Superfamily: ricin; rRNA N-glycosidase homology
C:Keywords: duplication; glycosidase; hydrolase; lectin; toxin
F:1-251/Product: abrin (clone 7.2) chain A #status predicted <ACH>
F:7-246/Domain: rRNA N-glycosidase homology <RNG>
F:74,113,195,196/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted
F:164,167/Active site: Glu, Arg #status predicted

alignment_scores:
Quality: 220.00 Length: 135
Ratio: 2.316 Gaps: 2
Percent Similarity: 70.370 Percent Identity: 40.741

alignment_block:

US-09-627-165B-15 x C39761 ..

Align seg 1/1 to: C39761 from: 1 to: 251

1 TACACAGATCTGGAGCGATACGCGGTCAT...AGGACCAAGATCCCTCT 47
|||||
113 TyrGlyAlaLeuGluArgTyrAlaHisGlnThrArgGluGlnIleSerLe 129
48 GGGTATAGGAGCAAGCATTCATCGCGCTCGCGCTGCTATTCACGACG 97
|||||
129 uGlyLeuAlaLeuThrHisAlaIleSerPheLeuArgSerIleValAs 146
98 GCAGACCCGGGCGCAAGCTGTTCCCTTATATCTCATTCAGATGATC 147
: : : : :
146 eAsnAspGluGluIuIysAlaArgThrLeuIleValIleIleGlnMeta 162
148 TCCGAGCGCGGAGATTCATCCCATCTTTGGAGGGCTCGCCATACAT 197
|||||
163 SerGluAlaAlaArgTyrArgTyrIleSerAsnArgValGlyValSer 179
198 TACAGCGGAGAGATCTTCCCGACATGATACATGCTCGAGCTGAGAGA 247
: : : : :
179 eArgThrGlyThrAlaPheGlnProAspProAlaMetLeuSerLeuGlu 196

248 CTAGTTGGGCGCCAGACATCCAGCAAGTCCAGACAGTACGATGCGCTT 297
: : : : :
196 snAsnTrpAspAsnLeuSerArgGlyValGlnIuIuSerValGlnAspHr 212
298 TTTTAATACCCATTTTCGTTG...CGTATATCCACCGGTAACCTTGTA 344
|||||
213 PheProAsnAlaValThrLeuArgArgValAsnAsnGlnProValIleVa 229
345 GTTTCAGCATGTTTCGGAGCTGATCGCCAGCTTACGATCATGTTGTTG 394
: |||
229 LaspSerLeuThrHisGlnSerValAlaValLeuAlaLeuMetLeuPhe 246
395 TATGT 399
|||||
246 aICys 247

seq_name: p1r2:B39761

seq_documentation_block:
abrin (clone 3.7) precursor - Indian licorice (fragment)
N:contains: rRNA N-glycosidase (EC 3.2.2.22)
C:Species: Abrus precatorius (Indian licorice)
C:Date: 21-Feb-1992 #sequence_revision 31-Dec-1993 #text_change 05-Jun-1998
C:Accession: B39761
R:Evensen, G.; Mathiesen, A.; Sundan, A.
J. Biol. Chem. 266: 6848-6852, 1991
A:Title: Direct molecular cloning and expression of two distinct abrin A-chains.
A:Reference number: A39761; MUID:91201329
A:Accession: B39761
A:Molecule type: DNA
A:Residues: 1-106 <EVE>
A:Cross-references: GB:X54873
C:Superfamily: ricin; rRNA N-glycosidase homology
C:Keywords: duplication; glycosidase; hydrolase; lectin; toxin
F:1-106/Product: abrin chain A (fragment) #status predicted <ACH>
F:1-101/Domain: rRNA N-glycosidase homology (fragment)
F:19,22/Active site: Glu, Arg #status predicted

alignment_scores:
Quality: 180.50 Length: 107
Ratio: 2.654 Gaps: 2
Percent Similarity: 63.551 Percent Identity: 41.121

alignment_block:

US-09-627-165B-15 x B39761 ..

Align seg 1/1 to: B39761 from: 1 to: 106

112 CAAGCTGTTCCCTTATATTCATTCATTCAGATGATCTCGAGCGCGAG 161
: : : : :
6 LysAlaArgThrLeuIleValIleIleGlnMetValSerGluAlaIleA 22
162 ATTCATCCCATCTTTGGAGGGCTCGCCATTCATTAACAGCGGGAGAT 211
|||||
22 gPheAlaGlyIleSerAsnArgValGlyValSerIleAlaGlyThrGlyTh 39
212 CATTTTCCTCCGACATGATACATCTCGAGCTGAGAGCTGATGGGCCAA 261
: : : : :
39 LaspGlnProAspProAlaMetIleSerLeuGlnAsnTrpAspAsn 55
262 CAATCCAGCAAGTCCAGCATGATGAGATGCGGTTTAAATACCATTT 311
|||||
56 LeuSerGlyGlyValGlnGlnSerValGlnGlyThrPheProHisPro 71
312 TCGGTTGGGTATATCCACCGGTAACCTTCGAGAGCTTGACGAAGT... 357
72ValThrLeuAlaGlnAsnIleCysA 79
358CGCAGCTGATCGCCACACTTA 378
79 snGluProValIleValAspSerLeuSerHisProThrValAlaValLeu 95

```

379 GCGATCATGTTGTTGTATGT 399
|||||
96 AlaLeuMetLeuPheValCys 102
seq_name: p1r1:RLTZR

seq_documentation_block:
  RNA N-glycosidase (EC 3.2.2.22) alpha-trichosanthin precursor [validated] - Mongolian S
  C:Keywords: abortifacient; glycosidase; hydrolase; root; toxin
  F:1-23/Domain: signal sequence #status predicted <SIG>
  F:24-270/Product: trichosanthin alpha #status experimental <MAT>
  F:27-266/Domain: RNA N-glycosidase homology <RNS>
  F:271-289/Domain: carboxyl-terminal propeptide #status predicted <CRP>
  F:93,183,186/Active site: Tyr, Glu, Arg #status predicted
  C:Date: 30-Sep-1988 #sequence_revision 26-Jan-1996 #text_change 23-Mar-2001
  C:Accession: J00566; A36274; J01093; A36273; J00003
  R:Shaw, P.C.; Yung, W.H.; Zhu, R.H.; Ho, W.K.K.; Ng, T.B.; Yeung, H.W.
  Gene 97, 267-272, 1991
  A:Title: Cloning of trichosanthin cDNA and its expression in Escherichia coli.
  A:Reference number: J00566; MUID:91153657
  A:Accession: J00566
  A:Molecule type: mRNA
  A:Residues: 1-289 <SMA>
  A:Cross-references: GB:M34858; NID:g170536; PIDN:AAA34207.1; PID:g170537
  R:Experimental source: tuber
  R:Chow, T.P.; Feldman, R.A.; Lovett, M.; Platak, M.
  J. Biol. Chem. 265, 8670-8674, 1990
  A:Title: Isolation and DNA sequencing of a gene encoding alpha-trichosanthin, a type I rib
  A:Reference number: A36274; MUID:90256790
  A:Accession: A36274
  A:Molecule type: DNA
  A:Residues: 1-233, 'V', 235-246, 'M', 248-289 <CHO>
  A:Cross-references: GB:J05434; NID:g170534; PIDN:AAA34206.1; PID:g170535
  R:Zheng, H.G.; Wang, P.; Shao, P.Z.; Yang, X.R.
  Acta Genet. Sin. 21, 42-51, 1994
  A:Title: Cloning and DNA sequencing of the gene encoding Trichosanthin.
  A:Reference number: J01093; MUID:94271613
  A:Accession: J01093
  A:Molecule type: DNA
  A:Residues: 1-72, 'V', 74-90, 'S', 92-233, 'T', 235-267, 'D', 269-289 <ZHE>
  A:Cross-references: GB:S70176; NID:g547148; PIDN:AB31048.1; PID:g547149
  R:Collins, E.J.; Robertus, J.D.; Lofrest, M.; Stone, K.L.; Williams, K.R.; Wu, P.; Hwan
  J. Biol. Chem. 265, 8665-8669, 1990
  A:Title: Primary amino acid sequence of alpha-trichosanthin and molecular models for ab
  A:Reference number: A36273; MUID:90256789
  A:Accession: A36273
  A:Molecule type: protein
  A:Residues: 24-270 <COL>
  R:Wang, Y.; Qian, R.Q.; Gu, Z.W.; Jin, S.W.; Zhang, L.Q.; Xia, Z.X.; Tian, G.Y.; Ni, C.Z
  Pure Appl. Chem. 58, 789-798, 1986
  A:Title: Scientific evaluation of Tian Hua Fen (THF): history, chemistry and application
  A:Reference number: J00003
  A:Accession: J00003
  A:Molecule type: protein
  A:Residues: 24-56, 'L', 58-59, 'T', 61-71, 'T', 73-81, 85-86, 'L', 88-92, 'DAGLEPRNAVL', 93-142, 'GL'
  R:Experimental source: tuber
  R:Huang, Q.; Liu, S.; Tang, Y.; Jin, S.; Wang, Y.
  Submitted to the Brookhaven Protein Data Bank, July 1994
  A:Reference number: A67091; PDB:IMR1
  A:Contents: annotation; X-ray crystallography, 1.6 angstroms, with adenine, residues 24-
  R:Huang, Q.; Liu, S.; Tang, Y.; Jin, S.; Wang, Y.
  Submitted to the Brookhaven Protein Data Bank, July 1994
  A:Reference number: A67092; PDB:IMR2
  A:Contents: annotation; X-ray crystallography, 1.6 angstroms, with formycin, residues 24
  R:Xiong, J.P.; Xia, Z.X.; Wang, Y.
  submitted to the Brookhaven Protein Data Bank, December 1994
  A:Reference number: A66711; PDB:ITCS
  A:Contents: annotation; X-ray crystallography, 1.7 angstroms, with MADPH, residues 24-27
  R:Xiong, J.P.; Xia, Z.X.; Wang, Y.
  Nat. Struct. Biol. 1, 695-700, 1994
  A:Title: Crystal structure of trichosanthin-NADPH complex at 1.7 Angstroms resolution re
  A:Reference number: A58622; MUID:95360714
  A:Contents: annotation; X-ray crystallography, 1.7 angstroms
  C:Comment: Alpha-trichosanthin has been used to induce abortions.
  A:gene: tcs
  C:Function:

```

```

A:Description: hydrolyzes the N-glycosidic bond of a specific adenosine in 28S rRNA t
C:Superfamily: RNA N-glycosidase; RNA N-glycosidase homology
C:Keywords: abortifacient; glycosidase; hydrolase; root; toxin
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-270/Product: trichosanthin alpha #status experimental <MAT>
F:27-266/Domain: RNA N-glycosidase homology <RNS>
F:271-289/Domain: carboxyl-terminal propeptide #status predicted <CRP>
F:93,183,186/Active site: Tyr, Glu, Arg #status predicted

alignment_scores:
  quality: 163.50      length: 170
  ratio: 1.587         gaps: 5
  percent similarity: 60.588      percent identity: 30.000

alignment_block:
  us-09-627-165b-15 x RLTZR ..
  Align seg 1/1 to: RLTZR from: 1 to: 289

1 TACACAGATCGGAGCGATACGCCGTCAT...AGGACCAAGATCCCTCT 47
|||||
134 TGTGluArgLeuGlnThrAlaIleAlaGlyLysIleArgGluAsnIleProle 150
|||||
48 GGGTATAGAGGAATCATTCATCCGTCGCGGCTTCGTTATCCAGGCG 97
|||||
150 uGlyLeuProAlaLeuAspSerAlaIleThrThrLeuPheTyrTyrAsnA 167
|||||
98 GCAGCACCCGCCGCCAAGCTCGTCCCTTAATTCCTCATTCAGATGATC 147
|||||
167 IAsnSerAlaIleAlaSerAla.....LeuMetValLeuIleGlnSerTrn 181
|||||
148 TCCGAGCCGCCGAGATTCATCCATCCATCTTTGGAGGCGTCGCCAATACAT 197
|||||
182 SerGluAlaIleAlaArgTyrLysPheIleGluGlnIleIleGlyLysArg 198
|||||
198 TAAACAGCGGGAGCTATTCCTCCGACATGATGATGATGATGATGATGATG 247
|||||
198 IAsp.....LysThrPheLeuProSerLeuAlaIleIleSerLeuGlnA 213
|||||
248 CTGTTGGCGGCCCAACATCCAGCAAGTCAG.....CAGTCTACGAT 291
|||||
213 snSerTrpSerAlaLeuSerLysGlnIleGlnIleAlaIleSerThrAsn 229
|||||
292 GCGGTTTTTAATACCATTCGTTGGTGGATATACCGGTAACCTTCCT 341
|||||
230 GlyGlnPheGluSerProValValLeuIleAsnAlaGlnAsnIleArgVa 246
|||||
342 GAGCTTGAGCAATGTTCCGACGCTGATCCGACGCTTAGCGATCATGTTGT 391
|||||
246 IThrIleThrAsnValAspAlaGlyValAlaThrSerAsnIleAlaLeuL 263
|||||
392 TTGTATGATAGGACGACCATCTCTCGACGTCGGTGGCTATTCGGCGCTG 441
|||||
263 euleuAsnArgAsnAsnMetAlaIleMetAsp..... 273
|||||
442 GTTCATGACGACCCGCTTGGAAAAATAGCGCGCGTCGACGATGTTACCTG 491
|||||
274 ..... 278
|||||
492 CACTGCTTC 501
|||||
278 tThrGlnSer 281
|||||

seq_name: p1r2:JC5606

seq_documentation_block:
  karasurin C - Trichosanthes kirilowii var. japonica
  N:Contains: karasurin A
  C:Species: trichosanthes kirilowii var. japonica
  C:date: 23-Sep-1997 #sequence_revision 23-Sep-1997 #text_change 20-Jun-2000
  C:Accession: J05606; JC5033
  R:Mizukami, H.; Iida, K.; Kondo, T.; Ogihara, Y.

```

Biol. Pharm. Bull. 20, 711-713, 1997
A:Title: Cloning and bacterial expression of a gene encoding ribosome-inactivating prote
A:Reference number: JCS606; MUID:97356562
A:Accession: JCS606
A:Molecule type: DNA
A:Residues: 1-289 <MIZ>
A:Cross-references: DDBJ:AB000666; NID:92329830; PID:BA21786.1; PID:92329831
R:Kondo, T.; Mizukami, H.; Takeda, T.; Ogihara, Y.
Biol. Pharm. Bull. 19, 1485-1489, 1996
A:Title: Amino acid sequences and ribosome-inactivating activities of karasurin-B and ka
A:Reference number: JCS032; MUID:97108848
A:Accession: JCS033
A>Status: preliminary
A:Molecule type: protein
A:Residues: 22-270 <KON>
C:Comment: This protein is a ribosome-inactivating protein and exhibit cytotoxic, aborti
C:Comment: This protein belongs to type I ribosomal-inactivating proteins which catalytl
C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology
C:Keywords: glycosidase
F:22-270/Product: karasurin C #status predicted <MNC>
F:24-270/Product: karasurin A #status predicted <MAA>
F:27-266/Domain: rRNA N-glycosidase homology <RNG>

alignment_scores:
Quality: 157.00 Length: 164
Ratio: 1.570 Gaps: 5
Percent Similarity: 60.976 Percent Identity: 29.878

alignment_block:
US-09-627-165B-15 x JCS606 ..

Align seg 1/1 to: JCS606 from: 1 to: 289

```
1 TACACAGTCGTGAGACGATACGCGGTGAT...AGGACGACGATCCCTCT 47
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
134 TyrgluatrgleuInlelaalaIaglyLysleargluAenilleProle 150
GGGATATAGAGACGATCAATCCGCTCGCGCTTCGTTATCCAGCG 97
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
150 uGlyleuProAlaLeuAspSerAlaIleThrThleuPheTrpYrAsnA 167
GCACGACCGCGGCGGAGCTGCTTCCCTTAAATCTCATTCAGATGATC 147
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
167 laaSenSerAlaIaSerAla.....leuMetValleuIleGlnSerThr 181
TCCGAGCGCGGAGATTCATCCATCTTTGGAGCGCTGCCAATACAT 197
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
182 SerIuAlaIaArgYrLysPheIleGluGlnIleGlyLysArgVa 198
198 TAAACGCGGCGGAGTCATTTCTCCGACATGTACATGCTGAGCTGAGA 247
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
198 lAsp.....LysThrPheLeuProSerleuAlaIleIleSerleuGluA 213
CTACTTGGGCGCAACAATCCAGCAGATCCAG.....CAGTCACGAT 291
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
213 snSerTyrSerAlaLeuSerLysGlnIleGlnIleAlaSerThrAsn 229
GGCGTTTAAATACCATTTCCGTTGGGTATATCCACCGTAACTTCGT 341
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
230 GlyInPheGluTrpProValIleuIleAsnAlaIleAsnGlnArgVa 246
GACGTTGAGCAATGTTCCGACGATGATCCGACGTTAGCATCATGTTGT 391
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
246 lThrIleThrAsnValAspAlaGlyValAlaThrSerAsnIleAlaLeu 263
392 TTGTATGTAGGAGCCGACCATCTTCTCCGACGTGCGTATTTGCGCGT 441
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
263 euLeuAsnArgAsn..... 267
442 GTCATACGACCGCTCTTGAAATATAGCGCGCGTGCAGAT 483
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
268 .....AsnMetAlaAlaIleAsp 274
```

seq_name: plr2:JE0401

seq_documentation_block:
antiviral protein - Virginian pokeweed
C:Species: Phytolacca americana (Virginian pokeweed)
C:Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 07-May-1999
C:Accession: JE0401
R:Kung, S.S.; Kimura, M.; Funatsu, G.
Agric. Biol. Chem. 54, 3301-3318, 1990
A:Title: The complete amino acid sequence of antiviral protein from the seeds of poke
A:Reference number: JE0401; MUID:91242096
A:Accession: JE0401
A:Molecule type: protein
A:Residues: 1-261 <KUN>
A:Experimental source: seed
C:Comment: This protein prevents the replication of a number of plant viruses, and in
C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology
C:Keywords: disulfide bond; glycoprotein
F:6-254/Domain: rRNA N-glycosidase homology <RNG>
F:10,44,255/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:34-258,84-105/Disulfide bonds: #status experimental

alignment_scores:
Quality: 156.50 Length: 135
Ratio: 1.720 Gaps: 3
Percent Similarity: 67.407 Percent Identity: 30.370

alignment_block:
US-09-627-165B-15 x JE0401 ..

Align seg 1/1 to: JE0401 from: 1 to: 261

```
1 TACACAGATCTGGAGGATACGCGGT.....CATAGGACGACATGCC 44
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
122 TyrProThrleuGluLysLysAlaGlyValThrSerArgAsnIleVal 138
TCTGGGTATAGAGAACTCAATCCGCTCGCGCTTCGTTATCCAG 94
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
138 nleuGlyIleGlnIleleuSerSerAspIleGlyLysIleSerGlyGln 155
GCGGCGACGACC...CGGCGCAAGCTCGTCCCTTAAATCTCATTCAG 141
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
155 LysSerPheThrGluLysIleGluAlaLysPheLeuValAlaIleGln 171
ATGATCTCCGAGCGCGGAGATTCATCCATCCATCTTTGAGGCGTCGCCA 191
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
172 MetValSerGluAlaAlaArgPheLysTyrIleGluAsnGlnValLys.. 187
192 ATACATTAACAGCGGAGTCATTTCTCCGACATGTACATGCTCGAGC 241
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
188 ....ThrAsnPheAsnArgAspPheSerProAsnAspLysValLeuAspL 203
TSGAGACTGTTGGGCGCAACAATCCAGCAGATCCAGAGTCATTCAGAT 291
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
203 euGluLysAsnTrpGlyLysIleSerThrAlaIleHisAsnSerLysAsn 219
GGCGTTTAAATACCATTTCCGTTGGGTATATCCACCGTAACTTCGT 341
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
220 GlyAlaLeuProLysProleuGluLeuLysAsnAlaAspGlyThrLysTr 236
GACGTTGAGCAATGTTCCGACGATGATCCGACGTTAGCATCATGTTGT 391
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
236 plleValleuArgValAspGluIleLysProAspAlaGlyLeuLeuAsnT 253
392 TTGTA 396
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
253 yTrVal 254
```

seq_name: plr2:S28421

seq_documentation_block:

OM of: US-09-627-165B-15 to: SwissProt_40:* out_format : pfs

About: Results were produced by the GenCore software, version 4.5
Copyright (c) 1993-2000 CompuGen Ltd.

```

Command line parameters:
-MODEL=framet-n2p.model
-O=cg92.1/USTO.spool/US09627765/runat.30072002-151140-3327/app_query.fasta_1.1012
-DB-Suffix=spool.40 -QEMT=fastan-Suffix=isp -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXP=0.000
-OGAPOP=4.500 -OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FCGAPOP=6.000 -FCGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=pl0sum62
-TRANS=human40.cdi -LIST=45 -DOCCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=fts
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USR=US09627765_cg92n1_1_69 -NCPU=6 -ICPU=3 -LOG=LOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPLY -WAIT -THREADS=1

```

```
Search information block:
Query: US-09-627-165B-15
Query length: 522
Database: Swissprot_40:*
Database sequences: 105224
Database length: 38719550
Search time (sec): 36.450000
```

Sequence	Strid	Orig	ZScore	EScore	Len	Document
SwissProt_40: M1A_VISAL +	621.00	1149.13	1.0e-56	254	1	P81446 vlsacm album (neurop
SwissProt_40: ABRR_ABRPR +	283.00	509.17	2.2e-21	528	1	P11140 abrus precatorius (in
SwissProt_40: R1CT_RICCO +	259.50	484.31	6.3e-19	576	1	P02879 ricinus communis (cas
SwissProt_40: ABRR_ABRPR +	256.00	458.01	1.5e-18	562	1	P28550 abrus precatorius (in
SwissProt_40: AGCL_RICCO +	256.00	447.97	1.5e-18	564	1	P06750 ricinus communis (cas
SwissProt_40: ABRR_ABRPR +	246.50	440.87	1.4e-17	527	1	P06077 abrus precatorius (in
SwissProt_40: NIGB_SAMNT +	186.00	326.98	2.9e-11	553	1	P33183 sambucus nigra (neuro
SwissProt_40: R1P2_BRKDI +	166.50	297.44	2.5e-09	282	1	P98184 bryonia dioica (red h
SwissProt_40: R1P2_BRKDI +	163.50	291.58	5.3e-09	289	1	P09989 trichosanthes kirilow
SwissProt_40: R1P2_TRIKI +	157.00	279.42	2.5e-08	289	1	P24478 phytoactaea americana
SwissProt_40: R1P5_PTKRI +	156.00	279.51	2.7e-08	261	1	P23339 phytoactaea americana
SwissProt_40: R1P4_POMXH +	156.00	277.37	3.2e-08	294	1	P03044 phytoactaea americana
SwissProt_40: R1P1_POMXH +	154.50	274.84	4.5e-08	286	1	P16094 momordica charantia (
SwissProt_40: R1P1_BRKDI +	145.50	257.86	4.0e-07	290	1	P33185 bryonia dioica (red h
SwissProt_40: R1P1_BRKDI +	144.50	255.22	5.1e-07	313	1	P10297 phytoactaea americana
SwissProt_40: R1P1_POMXH +	143.50	254.26	6.4e-07	286	1	P09194 cucumis figareli. puta
SwissProt_40: R1P1_POMXH +	141.50	250.24	1.0e-06	294	1	P56626 trichosanthes anguina
SwissProt_40: R1P1_TRAN +	141.00	249.58	1.2e-06	286	1	P29339 momordica balsamina (
SwissProt_40: R1P2_MONBA +	141.00	247.99	1.5e-06	278	1	P21326 mirabilis jalapa (gar
SwissProt_40: R1P2_MONBA +	137.00	241.09	3.1e-06	316	1	P33185 gellanum multiflorum
SwissProt_40: R1P4_GELMU +	134.50	237.73	5.5e-06	277	1	P00465 luffa cylindrica (smoc
SwissProt_40: R1P4_LUCY +	129.00	226.67	2.1e-05	299	1	P20656 saponaria officinalis
SwissProt_40: R1P5_SAOFF +	125.00	220.87	5.2e-05	253	1	P04389 saponaria officinalis
SwissProt_40: R1P5_SAOFF +	124.00	217.55	6.9e-05	292	1	P27559 saponaria officinalis
SwissProt_40: R1P7_SAOFF +	123.00	217.12	8.4e-05	250	1	P04391 saponaria officinalis
SwissProt_40: R1P7_SAOFF +	114.00	200.40	0.0007	250	1	P22851 luffa cylindrica (smoc
SwissProt_40: R1P8_LUCY +	112.00	195.06	0.0012	293	1	P24476 dianthus caryophyllus
SwissProt_40: R1P0_DIOCA +	103.50	181.33	0.0089	236	1	P27560 saponaria officinalis
SwissProt_40: R1P3_SAOFF +	93.50	159.58	0.1073	319	1	P09385 bacteriophage 933W. S
SwissProt_40: R1P2_PPMX +	91.00	155.19	0.1940	310	1	P04072 phytoactaea americana
SwissProt_40: R1P2_PPMX +	84.50	142.86	0.9277	315	1	P08082 bacteriophage h19b. S
SwissProt_40: S1TA_BPH19 +	84.50	142.86	0.9277	315	1	P10149 bacteriophage h10. sh
SwissProt_40: YN65_YEAST +	84.50	132.53	1.225	879	1	P42837 saccharomyces cerevisi
SwissProt_40: PML1_PRAV +	82.50	130.83	1.91	717	1	P06496 pinus avium (cherry) A
SwissProt_40: PADL_YEAST +	81.50	133.26	2.14	468	1	P55290 saccharomyces cerevisi
SwissProt_40: R1P4_POPFR +	80.50	127.12	3.08	715	1	P45730 populus trichocarpa (d
SwissProt_40: R1P9C_YEAST +	80.00	126.48	3.44	694	1	P38335 saccharomyces cerevisi
SwissProt_40: PML3_PHAU +	80.00	126.25	3.46	710	1	P31943 phaseolus vulgaris (K) A
SwissProt_40: PTH_DEIRA +	79.50	137.35	3.69	885	1	P00603 homo sapiens (human) A
SwissProt_40: PTH_DEIRA +	79.50	137.35	276	215	1	P09173 delnoccocus radioauran

seq_name	Swissprot_40:MIA_VISAL
Swissprot_40:BIAR_CHTLM	78.50 127.91 4.36
Swissprot_40:BLAD_BOVIN	78.50 127.67 4.39
Swissprot_40:RIAL_SHEEP	78.50 127.67 4.39
Swissprot_40:PIIL_DMCCA	78.50 123.47 4.96
Swissprot_40:UJ34_HSV1	77.50 131.13 4.79
	275
	P10218
	herpes simplex vira
	09pk6
	chlamydia muridar
	09t56
	bos taurus (bovine
	09t57
	ovis aries (sheep)
	02865
	daucus carota (car
	708
	herpes simplex vira

```
seq_documentation_block:
ID      MLA_VISAL      STANDARD;      PRT;      254 AA
```

DT 15-DEC-1998 (rel. 37, Created)
 DT 15-DEC-1998 (rel. 37, Last sequence update)
 DT 01-MAR-2002 (rel. 41, Last annotation update)
 DE Beta-galactoside specific lectin I A chain (MLA) (rRNA N
 DE glycosylase) (EC 3.2.2.42).
 OS Viscum album (European mistletoe).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 CC Santalales; Viscaceae; Viscum.
 ON NCBI_TaxID=3972;

RC STRAIN-SUBSP. ALBUM; MEDLINE-97134581; PubMed-6980141;
RX Huguet Soler M., Stoeva S., Schwaborn C., Wilhelm S., Stiefel T.,
RA Voelter W.;
RT "Complete amino acid sequence of the A chain of mistletoe lectin I."
RL FEBS Lett. 399:153-157(1996).
RE HUGUET SOLER M., STOEVA S., SCHWABORN C., WILHELM S., STIEFEL T., VOELTER W.
RW COMPLETE AMINO ACID SEQUENCE OF THE A CHAIN OF MISTLETOE LECTIN I.

CC - FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INITIATING PROTEIN
CC SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL
CC SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA. THE
CC B CHAIN BINDS TO CELL RECEPTORS AND PROBABLY FACILITATES THE ENTRY
CC INTO THE CELL OF THE A CHAIN. B CHAINS ARE ALSO RESPONSIBLE FOR
CC CELL AGGLUTINATION (LECTIN ACTIVITY).
CC - CATALYTIC ACTIVITY: Endoglycosidolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC - SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.
CC - MISCELLANEOUS: TWO ISOFORMS OF MLA EXIST: GLYCOSYLATED FORM MLA
CC AND NON-GLYCOSYLATED FORM MLA'.
CC - SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY. TYPE 2

DR InterPro; IPR001574; RIP

```
DR pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR PROSITE; PS00275; SHIGA_RICIN; FALSE_NEG.
```

KW	Hydrolase; Toxin; Repeat; Glycoprotein; Lectin.
FT	165 165 BY SIMILARITY.
FT	112 112 N-LINKED (GLCNAC. . .)
FT	CARBOHYD

	15	15	E -> D (IN MLA').
FT	VARIANT	15	
FT	VARIANT	66	V -> I (IN MLA').
FT	VARIANT	112	N -> T (IN MLA').

ET	VARIANT	112	112	p -> t (IN MLA').
ET	VARIANT	116	116	p -> t (IN MLA').
ET	VARIANT	133	134	DO -> EE (IN MLA').
ET	VARIANT	140	140	DO -> S (IN MLA').

FT	VARIANT	140	140	1	-> S (IN MLA') .
FT	VARIANT	144	144	1	-> Y (IN MLA') .
FT	VARIANT	151	151	1	-> A (IN MLA') .

FT	VARIANT	179	179	1 -> D (IN MLA').
FT	VARIANT	184	184	A -> E (IN MLA').
FT	VARIANT	190	190	V -> M (IN MLA').

FT	VARIANT	218	218	I -> F (IN MLA').
FT	VARIANT	223	224	pp -> ST (IN MLA').
FT	VARIANT	231	231	T -> S (IN MLA').

FT	VARIANT	235	235	D -> S (IN MLA').
SQ	SEQUENCE	254 AA;	28478 MM;	53BAF98D3E0FFE67 CRC64;

[illegible]

alignment_block:

US-09-627-165b-15 x MLA_VISAL ..

Align seg 1/1 to: MLA_VISAL from: 1 to: 254

```

1 TACACGAGATTCGGAGGATACCGCGTCATAGAGACGAGATCCCTCTGGG 50
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
115 TYPProAspLeuGluArgTyrAlaGlyHisArgAspGlnIleProLeuGln 131
51 TATAGAGAACTCATTCATCCGTCGGCGCTTCGTATATCCAGGCGCA 100
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
131 ylleAspGlnIleuIleGlnSerValThrAlaLeuArgPheProGlyGlyS 148
101 GCACCGGGCCCAAGCTCGTCCCTTATATATCTCATTCAGATGATCTCC 150
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
148 eRThrArgThrGlnAlaArgSerIleLeuIleLeuIleGlnMetIleSer 164
151 GAGGGCGGAGATTCATCCCATCTTTGGAGGCGCTGCCCATCATTTAA 200
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
165 GluAlaAlaArgPheAsnProIleLeuThrParGlyTyrArgGlnTyrIleAs 181
201 CAGCGGGAGATCATTTCTCCCGACATGATGCTGAGCTGGAGACTA 250
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
181 nserGlyIleSerPheLeuProAspValTyrMetLeuGlnLeuGlnThrS 198
251 GTTGGGGCCCAACATCCAGCAAGTCCAGCAGTTCAGGATGGCTTTT 300
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
198 eRTPGGLyGlnGlnSerThrGlnValGlnHisSerThrAspGlyValPhe 214
301 AATAACCATTTGGTGGTGGTATATCCACGCGTCACTTCGTCAGCTTGA 350
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
215 AsnAsnProIleArgLeuAlaIleProProGlyAsnPheValThrLeuTh 231
351 CAATTTGGCAGCTGATCCGACCTTACGATCATGTTGTTGATGTA 400
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
231 AsnValAlaArgAspValIleAlaSerLeuAlaIleMetLeuPheValCysG 248
401 GGCAGCGACCATCTCCGCC 420
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
248 LylGlnArgProSerSerSer 254
seq_name: SwissProt_40:ABRA_ABRPR
seq_documentation_block:
ID ABRA_ABRPR STANDARD: PRT: 528 AA.
AC P1140: P28589.
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Abrin-a precursor [Contains: Abrin-a A chain (rRNA N-glycosidase)]
DE (EC 3.2.2.22): Abrin-a B chain.
OS Abrus precatorius (Indian licorice) (Crab's eye).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
OX NCBI_Taxid=3816;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93132798; PubMed=8421313.
RA Hung C.-H., Lee M.-C., Lee T.C., Lin J.-Y.;
RT "Primary structure of three distinct isoabins determined by cDNA
RT sequencing. Conservation and significance."
RL J. Mol. Biol. 229:263-267(1993).
RN [2]
RP SEQUENCE OF 1-251.
RC TISSUE=Seed;
RA Funatsu G., Taguchi Y., Kamenosono M., Yanaka M.;
RT "The complete amino acid sequence of the A-chain of abrin-a, a toxic
RT protein from the seeds of Abrus precatorius."
RL Agric. Biol. Chem. 52:1095-1097(1988).
RN [3]
RP SEQUENCE OF 1-251 FROM N.A.
RC TISSUE=Leaf;

```

```

RX MEDLINE=91201329; PubMed=2016300;
RA Evensen G., Mathiesen A., Sundan A.;
RT "Direct molecular cloning and expression of two distinct abrin
RT A-chains."
RL J. Biol. Chem. 266:6848-6852(1991).
RN [4]
RP SEQUENCE OF 262-528.
RX MEDLINE=92371656; PubMed=1505674;
RA Chen Y.-L., Chow L.-P., Tsugita A., Lin J.-Y.;
RT "The complete primary structure of abrin-a B chain."
RL FEBS Lett. 309:115-118(1992).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.14 ANGSTROMS).
RX MEDLINE=95333188; PubMed=7608980;
RA Tanirov T.H., Lu T.-H., Liaw Y.-C., Chen Y.-L., Lin J.-Y.;
RT "Crystal structure of abrin-a at 2.14 A."
RL J. Mol. Biol. 250:354-367(1995).
CC -1- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN
CC SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL
CC SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA.
CC ABRIN-A IS MORE TOXIC THAN RICIN.
CC -1- FUNCTION: THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT
CC FACILITATES THE BINDING OF ABRIN TO THE CELL MEMBRANE THAT
CC PRECEDES ENDOCYTOSIS.
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -1- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.
CC -1- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN
CC CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO THE RIBOSOME-
CC INACTIVATING FAMILY, TYPE 2 RIP SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 2 RICIN B-TYPE LECTIN DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: M98344; AAA32624.1; ALT_INT.
DR EMBL: X54872; -; NOT_ANNOTATED_CDS.
DR PIR: S2429; T2LSA.
DR PIR: S24133; S24133.
DR PDB: 1ABR; 07-FEB-95.
DR InterPro: IPR001574; RIP.
DR InterPro: IPR000772; Ricin_B_lectin.
DR Pfam: PF00652; Ricin_B_lectin; 2.
DR Pfam: PF00161; RIP; 1.
DR PRINTS: PR00396; SHIGARICIN.
DR SMART: SM00458; RICIN; 2.
DR PROSITE: PS00231; RICIN_B_LECTIN; 2.
DR PROSITE: PS00275; SHIGA_RICIN; 1.
KW Hydroxylase; Protein synthesis inhibitor; Toxin; Repeat; Glycoprotein;
KW Lectin; 3d-structure.
FT CHAIN 1 251
FT PEPTIDE 252 261 ABRIN-A A CHAIN.
FT CHAIN 262 528 LINKER PEPTIDE.
FT DOMAIN 273 400 ABRIN-A B CHAIN.
FT DOMAIN 403 527 RICIN B-TYPE LECTIN 1.
FT REPEAT 283 325 RICIN B-TYPE LECTIN 2.
FT REPEAT 326 366 1-ALPHA.
FT REPEAT 369 401 1-BETA.
FT REPEAT 414 449 1-GAMMA.
FT REPEAT 453 492 2-ALPHA.
FT REPEAT 495 528 2-BETA.
FT ACT_SITE 164 164 2-GAMMA.
FT DISULFID 247 269 BY SIMILARITY.
FT DISULFID 286 305 INTERCHAIN (BY SIMILARITY).
FT DISULFID 329 346 BY SIMILARITY.
FT DISULFID 417 430 BY SIMILARITY.
FT DISULFID 456 473 BY SIMILARITY.

```


FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT CARBOHYD 361 361 N-LINKED (GLCNAC. . .).
FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .).
FT CONFLICT 202 202 MISSING (IN REF. 2).
FT CONFLICT 238 238 M -> Y (IN REF. 4).
FT CONFLICT 427 427 M -> L (IN REF. 4).
FT CONFLICT 467 467 T -> P (IN REF. 4).
FT CONFLICT 483 483 V -> L (IN REF. 4).
SQ SEQUENCE 528 AA: 59243 MW: 11768BCD5B9A827 CRC64:

alignment_scores: Quality: 283.00 Length: 185
Ratio: 2.320 Gaps: 6
Percent Similarity: 65.946 Percent Identity: 38.919

alignment_block:

US-09-627-165B-15 x ABRA_ABRPR ..

Align seg 1/1 to: ABRA_ABRPR from: 1 to: 528

1 TCACAGATCTGGAGCATAGCGCGTCAT...AGGACGACATCCCTCT 47
|||||
113 TYGTGlyAspLeuGluArgTrpAlaHisGlnSerArgGlnGlnIleProLe 129
48 GGGTATAGAGAACTCATTCATCCGCTCGCGCTGCTTATCCAGCG 97
|||||
129 uGlyLeuGlnAlaLeuThrHisGlyIleSerPhePheArgSerGlyGly 146
98 GCGACGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 147
|||||
146 snAspAsnGlnGluIuysAlaArgThrLeuIleValIleIleIleMetVal 162
148 TCCGAGCGCGCGAGATTCATCCATCTTTGGAGGCGTCCCAATACAT 197
|||||
163 AlagluAlaAlaArgPheArgTrpIleSerAsnArgValArgValSer11 179
198 TAACAGCGCGGAGCATTTCTCCGACATGTACATCCGACGACGTCGAG 247
|||||
179 eGlnThrGlyThrAlaPheGlnProAspAlaAlaMetIleSerLeuGlnA 196
248 CTAGTGGGGGCGACAAATCCACGCAAGTCGAGAGTCTACGATGGCGTT 297
|||||
196 snAsnTrpAspAsnLeuSerArgGlyValGlnGlnIleSerValGlnAspTr 212
298 TTATATACCATTCGCTGGGTGGATATATCCACGCGTAACTTCGTGACGT 347
|||||
213 PheProAsnGln.....ValThrIle 219
348 GAGCAATGTCGCGAC.....G 364
|||||
219 uThrAsnIleArgAsnGlnProValIleValAspSerLeuSerHisProT 236
365 TGATGCGCAGCTTACGATCATGTTGTTGATATAGGACCGCGCAATCT 414
|||||
236 hValAlaIleValAlaLeuAlaMetLeuPheValCys...AspProProAsn 251
415 TCCTCGGAGCGTGGCTATTTGGCGGTGATACGACCGCGCTTGGAAGA 464
|||||
252 AlaAsnGlnSer.....ProLeuLeuIleArgSerIleValGlnIuLy 265
465 TAGCGCGCGCGCGTGCAGATGTATACGATGCTGCTCCGACCGCGCGC 514
|||||
265 sSer.....LysIleCysSerSerArgTrpGluProThrValA 278
515 GCATC 519
|||||
278 rgIle 279

seq_name: swissprot_40:RIC1_RICCO
seq_documentation_block:
ID RIC1_RICCO STANDARD: PRT: 576 AA.

AC P02879: P02880:
DT 21-JUL-1986 (rel. 01, Created)
DT 13-AUG-1987 (rel. 05, Last sequence update)
DT 01-MAR-2002 (rel. 41, Last annotation update)
DE Ricin precursor (contains: Ricin A chain (rRNA N-glycosidase)
DE (EC 3.2.2.22); Ricin B chain).
OS Ricinus communis (Castor bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eudicots I; Malpighiales; Euphorbiaceae; Ricinus.
OX NCBI_TaxID=3986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86067214; PubMed=2999712;
RA Halling K.C., Halling A.C., Murray E.E., Ladin B.F., Houston L.L.,
RA Weaver R.F.,
RT "Genomic cloning and characterization of a ricin gene from Ricinus
RT communis."
RL Nucleic Acids Res. 13:8019-8033(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92163016; PubMed=1371405;
RA Tregear J.W., Roberts L.M.,
RT "The lectin gene family of Ricinus communis: cloning of a functional
RT ricin gene and three lectin pseudogenes."
RL Plant Mol. Biol. 18:515-525(1992).
RN [3]
RP SEQUENCE OF 12-576 FROM N.A.
RX MEDLINE=85179479; PubMed=3838723;
RA Lamb A., Roberts L.M., Lord J.M.,
RT "Nucleotide sequence of cloned cDNA coding for preproricin."
RL Eur. J. Biochem. 148:265-270(1985).
RN [4]
RP SEQUENCE OF 36-302.
RA Yoshitake S., Funatsu G., Funatsu M.,
RT "Isolation and sequences of peptic peptides, and the complete
RT sequence of the chain of ricin-D."
RL Agric. Biol. Chem. 42:1267-1274(1978).
RN [5]
RP SEQUENCE OF 315-576.
RA Funatsu G., Kimura M., Funatsu M.,
RT "Primary structure of Ala chain of ricin D."
RL Agric. Biol. Chem. 43:2221-2224(1979).
RN [6]
RP CARBOHYDRATE-LINKAGE SITES, AND PARTIAL SEQUENCE.
RX MEDLINE=90344223; PubMed=1368517;
RA Kimura Y., Kusoku H., Tada M., Takagi S., Funatsu G.,
RT "Structural analyses of sugar chains from ricin A-chain variant."
RL Agric. Biol. Chem. 54:157-162(1990).
RN [7]
RP REVIEW.
RX MEDLINE=21480122; PubMed=11595634;
RA Olshes S., Kozlov J.V.,
RT "Ricin."
RL Toxicon 39:1723-1728(2001).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
RX MEDLINE=87165983; PubMed=3558397;
RA Monfort W., Villafraña J.E., Monzingo A.F., Ernst S.R., Katzin B.,
RT "The three-dimensional structure of ricin at 2.8 A."
RL J. Biol. Chem. 262:5396-5403(1987).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF A-CHAIN.
RX MEDLINE=91352004; PubMed=1881881;
RA Katzin B.J., Collins E.J., Robertus J.D.,
RT "Structure of ricin A-chain at 2.5 A."
RL Proteins 10:251-259(1991).
RN [10]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF B-CHAIN.
RX MEDLINE=91352005; PubMed=1881882;
RA Rutenber E., Robertus J.D.,
RT "Structure of ricin B-chain at 2.5-A resolution."

RL Proteins 10:260-269(1991).
 RL [11]
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF A-CHAIN.
 RX MEDLINE-95082010; PubMed-7990130.
 RA Weston S.A., Tucker A.D., Thatcher D.R., Dedyshire D.J.,
 RT Paupit R.A.;
 RT "X-ray structure of recombinant ricin A-chain at 1.8-A resolution.";
 RN J. Mol. Biol. 244:410-422(1994).
 RN [12]
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF MUTANT HIS-215.
 RX MEDLINE-96374222; PubMed-8780513;
 RA Day P., Ernst S.R., Frankel A.E., Monzingo A.F., Pascal J.M.,
 RT Molina-Svane M.C., Robertus J.D.;
 RT "Structure and activity of an active site substitution of ricin A
 chain";
 RL Biochemistry 35:11098-11103(1996).
 RL [13]
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF A-CHAIN.
 RX MEDLINE-97240820; PubMed-9086280;
 RA Yan X., Hollis T., Svane M., Day P., Monzingo A.F., Milne G.W.,
 RT Robertus J.D.;
 RT "Structure-based identification of a ricin inhibitor";
 RN J. Mol. Biol. 266:1043-1049(1997).
 RN [14]
 RP MUTAGENESIS.
 RX MEDLINE-93165632; PubMed-1287657;
 RA Kin Y., Robertus J.D.;
 RT "Analysis of several key active site residues of ricin A chain by
 mutagenesis and X-ray crystallography.";
 RL Protein Eng. 5:775-779(1992).
 RL -1- FUNCTION: Ricin is highly toxic to animal cells and to a less
 extent to plant cells. The A chain is responsible for inhibiting
 protein synthesis through the catalytic inactivation of 60S
 ribosomal subunits. It act as a glycosylase that removes that a
 specific adenine residue from an exposed loop of 28S ribosomal
 RNA. As this loop is involved in the binding of elongation
 factors, the modified ribosomes are unable to support protein
 synthesis. The A chain can inactivate a few thousands ribosomes
 per minute, thus inactivating them faster than the cell can make
 new ones. A single A-chain molecule can therefore kill an animal
 cell. The B chain binds to cell receptors and facilitates the
 entry into the cell of the A chain. B chains are also responsible
 for cell agglutination (lectin activity). It binds to beta-D-
 galactopyranoside moieties.
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
 specific adenosine on the 28S rRNA.
 CC -1- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.
 CC -1- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN
 CC CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).
 CC -1- PTM: THE MAJOR A-CHAIN IS GLYCOSYLATED ONLY IN POSITION 45 AND THE
 CC MINOR A-CHAIN IS GLYCOSYLATED ALSO IN POSITION 271.
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO THE RIBOSOME-
 CC INACTIVATING FAMILY. TYPE 2 RIP SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 2 RICIN B-TYPE LECTIN DOMAINS.
 CC -1- CAUTION: REF.4 AND REF.5 SEQUENCES HAVE A HIGH NUMBER OF CONFLICTS
 CC WITH THE SEQUENCE TRANSLATED FROM DNA (REF.1, REF.2 AND REF.3).
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
 CC The European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X031179; CAA26939.1; -
 DR EMBL; X52908; CAA37095.1; -
 DR EMBL; X02388; CAA36230.1; -
 DR EMBL; A12892; CAA01058.1; -
 DR PIR; A24041; RLCSD.
 DR PDB; 2AAT; 31-JAN-94.
 DR PDB; IAPG; 31-JAN-94.
 DR PDB; IFMP; 31-OCT-93.
 DR

	PDB; 1JFS; 14-JAN-98.	DR
	PDB; 1JFY; 14-JUN-98.	DR
	PDB; 1IFU; 14-JUN-98.	DR
	PDB; 1RTC; 31-OCT-93.	DR
	PDB; 1OBS; 16-JUN-97.	DR
	PDB; 1OBT; 16-JUN-97.	DR
	PDB; 1BR5; 02-SEP-98.	DR
	PDB; 1BR6; 02-SEP-98.	DR
	GlycosultedB; P02879; -.	DR
	InterPro; IPR001574; R.P.	DR
	InterPro; IPR000772; Ricin_B_lectin.	DR
	pfam; PF00652; Ricin_B_lectin. 2.	DR
	pfaam; pf00161; RIP; 1.	DR
	PRINTS; PR00396; SHIGARICIN.	DR
	SMART; SMO0458; RICIN; 2.	DR
	PROSITE; PS50231; RICIN_B_LECTIN; 2.	DR
	PROSITE; PS00275; SHIGA_RICIN; 1.	KW
	Hydrolase; Protein synthesis inhibitor; Toxin; Repeat; Glycoprotein;	KM
	Lectin; signal; 3D-structure.	
FT SIGNAL	1 35	
FT CHAIN	36 302	RICIN A CHAIN.
FT PEPTIDE	303 314	LINKER PEPTIDE.
FT CHAIN	315 576	RICIN B CHAIN.
FT DOMAIN	321 448	RICIN B-TYPE LECTIN 1.
FT DOMAIN	451 575	RICIN B-TYPE LECTIN 2.
FT REPEAT	331 373	1-ALPHA.
FT REPEAT	374 414	1-BETA.
FT REPEAT	417 449	1-GAMMA.
FT REPEAT	462 497	2-ALPHA.
FT REPEAT	501 540	2-BETA.
FT REPEAT	543 570	2-GAMMA.
FT ACT_SITE	212 212	
FT DISULFID	294 318	INTERCHAIN.
FT DISULFID	334 353	
FT DISULFID	377 394	
FT DISULFID	465 478	
FT CARBOHYD	504 521	N-LINKED (GLCNAC. .).
FT FT	45 45	/FTid-CAR..000080.
FT FT	271 271	N-LINKED (GLCNAAC. .); IN MINOR A-CHAIN VARIANT.
FT CARBOHYD	409 409	/FTid-CAR..000081.
FT CARBOHYD	449 449	N-LINKED (GLCNAC. .)
FT CONFLICT	76 76	E -> D (IN REF. 3).
FT CONFLICT	551 551	A -> R (IN REF. 3).
FT STRAND	43 47	
FT TURN	49 50	
FT HELIX	53 67	
FT STRAND	73 74	
FT TURN	75 76	
FT STRAND	77 79	
FT TURN	88 90	

alignment_scores:

Quality:	259.50	Length:	178
Ratio:	2.257	Gaps:	5
Percent Similarity:	64.607	Percent Identity:	41.011

alignment_block:
US-09-627-165B-15 x RICL_RTCCO ..

```
Align seg 1/1 to: RICL_RTCCO from: 1 to: 576
```

```

      1 TACAGATGTGAGCGATGCCGGTCAT...AGGCAGATCCCTCT   47
        |||          |         :|||:::    |||:::||||   ||
     158 TyrsaphtyglengluInleuaIlaGlYaslnleatrglunasmllcglude 174
       GGGATTAGAGAACTCATTCATTCGTCGTCGGTGCTGTTA.....  90
        ||||  |||  ||| ::::::::::||| |||||  ||| 
     174 ugliysngilyProleugnlunglnalileserAlaleutyrytyrYrseET 191
           91 ..CAAGCGGAGACC GGCCCCAACAGCTCTGTCCTTAATAATCCGCA TT 138

```



```

197 MetIleSerGluAlaAlaArgPheGlnTyrIleGluGlyLumetaArgH 213
192 ATACATTACAGCGGGAGTCATTTCTCCGACATGATACATCTGGAG 241
213 FargIleArgTyrAsnArgSerAlaProAsnProSerValIleHrl 230
242 TGGAGACATAGTGGGGCCACATCCAGCAGACGATCTACGAT 291
230 eGluAsnSerTrpPolyArgLeuSerThrAlaIleGlnIleGlnSerAangln 246
292 GGCCTTTTAAATACCATTTGGTGGTATATCCACCGCTACTTCGT 341
247 GlyAlaIlePheAlaSerProIleGlnIleuAlaArgAsnGlySerLysph 263
342 GACGTTGACGATGTTCCGACGATGATCCAGCTACGATGTTGT 391
263 eAsnValTyrAspValSerIleLeuIleProIleIleAlaLeuMetValT 280
392 TTGTATGTAGGAGCGACCATCTTCCTCCGACGTCGCTATGGCCCTG 441
280 YrArgCysAlaProProProSerSerGln.....PheSerLeu 292
442 GTCATACGACCGCTCTTGGAAATAGCGCGCGCTCGACGATGTACTGT 491
293 LeuIleArgProValValProAsnPheAsnAla.....AspVal...Cy 306
492 CACTGCTCCGACCGACCGCTGCGCATCGTA 522
306 sMetAspProGluProIleValArgIleVal 316

seq_name: SwissProt_40:ABRB-ABRP
seq_documentation_block:
ID ABRB ABRP STANDARD: PRT: 527 AA.
AC 006077: P81374:
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE AbriIn-b precursor [Contains: AbriIn-b A chain (rRNA N-glycosidase)
DE (EC 3.2.2.22): AbriIn-b B chain].
OS Abrus precatorius (Indian licorice) (Crab's eye).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eufrosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
OX NCBI_TaxID=3816;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93132798; PubMed=8421313;
RA Hung C.-H., Lee M.-C., Lee T.C., Lin J.-Y.,
RT "Primary structure of three distinct isobritins determined by cDNA
RT sequencing. Conservation and significance.";
RL J. Mol. Biol. 229:263-267(1993).
RN [2]
RP SEQUENCE OF 260-527.
RC TISSUE=Seed; PubMed=7763422;
RX MEDLINE=93169023; PubMed=7763422;
RA Kimura M., Sumitawa T., Funatsu G.,
RT "The complete amino acid sequences of the B-chains of abrin-a and
RT abrin-b, toxic proteins from the seeds of Abrus precatorius.";
RL Biosci. Biotechnol. Blochem. 57:166-169(1993).
CC -1- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN
CC SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL
CC SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RRNA.
CC ABRIN-A IS MORE TOXIC THAN RIBIN.
CC -1- FUNCTION: THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT
CC FACILITATES THE BINDING OF ABRIN TO THE CELL MEMBRANE THAT
CC PRECEDES ENDOCYTOSIS.
CC -1- CATALYTIC ACTIVITY: Endohydrolase of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -1- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.
CC -1- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN
CC CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-

```

```

CC INACTIVATING FAMILY, TYPE 2 RIP SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 2 RICIN B-TYPE LECTIN DOMAINS.
-----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: M98345; AAA32625.1; -.
DR HSP: P11140; IABR.
DR InterPro: IPR001574; R1Cin_B_lectin.
DR InterPro: IPR000772; R1Cin_B_lectin.
DR Pfam: PF00652; R1Cin_B_lectin; 2.
DR Pfam: PF00161; R1P; 1.
DR PRINTS: PRO0396; SHIGARICIN.
DR SMART: SM00458; RICIN; 2.
DR PROSITE: PS50231; RICIN_B_LLECTIN; 2.
DR PROSITE: PS00275; SHIGA_RICIN; 1.
KW Hydrolase; Protein synthesis inhibitor; Toxin; Repeat; Glycoprotein;
KM Lactin.
FT CHAIN 1 250 ABRIN-B A CHAIN.
FT PEPTIDE 251 260 LINKER PEPTIDE.
FT CHAIN 261 527 ABRIN-B B CHAIN.
FT DOMAIN 272 399 RICIN B-TYPE LECTIN 1.
FT DOMAIN 402 526 RICIN B-TYPE LECTIN 2.
FT REPEAT 282 324 1-ALPHA.
FT REPEAT 325 365 1-BETA.
FT REPEAT 368 400 1-GAMMA.
FT REPEAT 413 448 2-ALPHA.
FT REPEAT 452 491 2-BETA.
FT REPEAT 494 527 2-GAMMA.
FT ACT_SITE 163 163 INTERCHAIN (BY SIMILARITY).
FT DISULFID 285 304 BY SIMILARITY.
FT DISULFID 328 345 BY SIMILARITY.
FT DISULFID 416 429 BY SIMILARITY.
FT DISULFID 455 472 PYRROLIDONE CARBOXYLIC ACID (BY
FT MOD_RES 1 1 SIMILARITY).
FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 360 360 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 400 400 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 282 282 N -> D (IN REF. 2).
FT CONFLICT 291 291 D -> N (IN REF. 2).
FT CONFLICT 350 351 AE -> PQ (IN REF. 2).
FT CONFLICT 378 378 S -> N (IN REF. 2).
FT CONFLICT 426 426 L -> M (IN REF. 2).
FT CONFLICT 428 428 Y -> D (IN REF. 2).
FT CONFLICT 431 431 N -> S (IN REF. 2).
FT CONFLICT 431 431 R -> K (IN REF. 2).
FT CONFLICT 484 484 N -> S (IN REF. 2).
FT CONFLICT 491 491 H -> Y (IN REF. 2).
FT CONFLICT 493 493 R -> G (IN REF. 2).
FT CONFLICT 502 502 E -> Q (IN REF. 2).
FT CONFLICT 509 509 H -> T (IN REF. 2).
FT CONFLICT 513 516 H -> T (IN REF. 2).
SQ SEQUENCE 527 AA; 59114 MW; 3253AE490CB9494A CRC64;

alignment_scores:
Quality: 246.50 Length: 175
Ratio: 2.071 Gaps: 6
Percent Similarity: 68.000 Percent Identity: 38.857

alignment_block:
US-09-627-165b-15 x ABRB-ABRP
Align seq 1/1 to: ABRB-ABRP from: 1 to: 527
1 TACACAGATCTGAGCGCATACGCGGTCAAT..AGGACACAGATCCCTCT 47

```

```

113 TyrILAspleuGlnAlaArgLeuAlaIaArgInThrArgInGlnIleProLe 129
148 GGGATATAGAGAACTCATTTCAATCCGCTGCGGCGTTCGTATCCAGCG 97
129 uGlyLeuGlnAlaLeuArgHisAlaIleSerPheLeuGln...SerGlyT 145
98 GCACACCCGGGCGCCAGCTGCTGCTCCCTATATACCTCATTTACAGATGC 147
145 hrAspArgInGlnIleAlaArgThrLeuIleValAlaIleIleIleMetAla 161
148 TCCGAGGCCCGCGAGATTCATCCATCCATCTTTTGAGAGCGCTCGCAATACAT 197
162 SerGlnAlaAlaArgTyrArgPheIleSerTyrArgValGlyValSerTl 178
198 TAACAGCGGGAGTCAATTTCTTCCGACATGTACATGCTCGAGCTGGAGA 247
178 eArgTThrAsnThrAlaPheGlnProAspAlaAlaMetIleSerLeuGlnAla 195
248 CTATGTTGGGGCCACAAATCCACGACCAAGCAAGCAGCATTCAGAGGGCGT 257
195 snAsnTyrAspAsnLeuSerGlyGlyValGlnGlnInSerValGlnAspThr 211
298 TTTAAATAACCAATTCGTTG...GATATATCCACCGGTAACTTCGTGAC 344
212 PheProAsnAlaValAlaThrLeuArgSerValAsnAsnGlnProValIleVal 228
345 GTTGAACAAATGTTCCGACGATGATCGACGCGATGTGAGCATGTTGTTG 394
228 IAspSerLeuThrHisGlnSerValAlaValAlaLeuAlaMetLeuPheV 245
395 TATGTAGGAGCGACCATCTTCCTCCCGACGTGGCGGTATGGCCGCTGATC 444
245 aATCys...AsnProProAsnAlaGlnInSer.....ProLeuLeu 257
445 ATAGACACCCGCTTGTGAATAATAGCGCGCGCTGACAGATGTATACCTGCAC 494
258 ILeArgSerIleValGlnIlySer.....LysIleSerSer 270
495 TGCTTCGACACCCAGCGCTGGCGCATC 519
270 rArgTyrGlnProThrValArgIle 278

seq_name: SwissProt_40:NIGB_SAMNI

seq_documentation_block:
ID NIGB_SAMNI STANDARD: PRF: 563 AA.
AC P33183; P33184; P93542;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Niglin b precursor (Agglutinin V) (SNAB) [Contains: Nigrin b A chain
  (rRNA N-glycosidase) (EC 3.2.2.22); Nigrin b B chain)].
OS Sambucus nigra (European elder).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  Asteridae; euasterids II; Adoxaceae; Sambucus.
NCBI_TaxID=4202;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=Bark;
RX MEDLINE=96215449; PubMed=8647092;
RA Van Damme E.J., Barre A., Rouge P., Van Leuven F., Peumans W.J.;
  "Characterization and molecular cloning of Sambucus nigra agglutinin V
  (nigrin b), a GalNAc-specific type-2 ribosome-inactivating protein
  from the bark of elderberry (Sambucus nigra).";
  Eur. J. Biochem. 237:505-513(1996).
RN [2]
RP SEQUENCE OF 26-49 AND 298-321.
RC TISSUE=Bark;
RX MEDLINE=94003077; PubMed=8400135;
RA Gibbs T., Cifuentes L., Ferreira J.M., Rojo M.A., Iglesias R.,
  Munoz R., Atlas F.J., Calonge M., Garcia J.R., Mendez E.;

```

RT	"Isolation and partial characterization of nigrin b, a non-toxic novel type 2 ribosome-inactivating protein from the bark of Sambucus nigra L.";
RU	Plant Mol. Biol.
CC	-1- FUNCTION: NON-TOXIC TYPE 2 RIP WHICH STRONGLY INHIBITS MAMMALIAN PROTEIN SYNTHESIS BUT DOES NOT AFFECT PLANT NOR BACTERIAL PROTEIN SYNTHESIS. THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL SUBUNIT BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA. THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT FACILITATES THE BINDING OF NIGRIN B TO THE CELL MEMBRANE THAT PRECEDES ENDOCYTOSIS.
CC	-1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.
CC	-1- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.
CC	-1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-INACTIVATING FAMILY, TYPE 2 RIP SUPERFAMILY.
CC	-1- SIMILARITY: CONTAINS 2 RICIN B-TYPE LECTIN DOMAINS.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. usage by and for commercial entities requires a license agreement (See http://www.isb-sdb.ch/announce/ or send an email to license@isb-sdb.ch).
CC	EMBL: U41299, AB836475.1, -. .
DR	PIR: S37382; S37382.
DR	PIR: S37383; S37383.
DR	HSSP: P02879; ZNAT.
DR	Mendel; 15504; Samml:R1P;15504.
DR	InterPro: IPR000772; Ricin_B_lectin.
DR	InterPro: IPR001574; RIP.
DR	Pfam: PF00652; Ricin_B_lectin; 2.
DR	Pfam: PF00161; RIP; 1.
DR	PRINTS: PR00396; SHIGARICIN.
DR	SMART: SM00458; RICIN; 2.
DR	PROSITE: PS00275; SHIGA_LECTIN; 1.
DR	PROSITE: PS50231; RICIN_B_LECTIN; 2.
KW	Hydrolase; Protein synthesis inhibitor; Toxin; Repeat; Glycoprotein; Lectin; Signal.
FT	SIGNAL
FT	CHAIN
FT	CHAIN
FT	DOMAIN
FT	DOMAIN
FT	REPEAT
FT	REPEAT
FT	REPEAT
FT	REPEAT
FT	REPEAT
FT	ACET_SITE
FT	DISULFID
FT	DISULFID
FT	DISULFID
FT	DISULFID
FT	CARBOHYD
FT	CARBOHYD
FT	CARBOHYD
FT	CARBOHYD
FT	CARBOHYD
FT	CONFLICT
SEQUENCE	563 AA; 62300 MW; F250CBB24621.P1A.Cp654.

```
alignment_scores:
  Quality: 186.00
  Ratio: 1.661
  Percent Similarity: 65.116
  Length: 172
  Gaps: 7
  Percent Identity: 33.721
```

alignment block:
US-09-627-165b-15 x NIGB_SAMNI ..

Align seg 1/1 to: NIGB_SAMNI from: 1 to: 563

```

1 TACACAGATCTGAGGATACGCCGCT...CATAGGACCAATCCCTCT 47
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
141 TyrapasnleugluThralaalaAsnThrArgArgGlnSerIleGluLe 157
48 GCGATAGAGGAAGATCATTCATCCGTCGGTCCGCTTCATTACAGGCC 97
157 uGlyProSerProleuaspelYalalIethrSerLeu...TyhIsGlyA 173
98 GCACACACCGGGCCCAAGCTGTTCCCTTAATATCTCATTCAGATGATC 147
173 spSerVal.....AlargSerleuValIalIegInMetVal 186
148 TCCGAGGCGCGGAGATTCATCCATCTTTGGAGGGCTCGCCATATCAT 197
187 SerGluAlaAlaArgPheArgYrIleGluGluGluValAlaArgArgSe 203
198 TAACAGCGGAGATCATTTCTTCCGACATGTACATGCTCAGCTGAGAGA 247
203 uGlnGlnAlaThrSerPheThrProAlaAlaLeuMetLeuSerMetGluA 220
248 CTAGTGGGCGCCAAATCCACGAGATCCAGAGCTCTAGGAGTGGCGGT 297
220 snAsnTrpSerSerMetSerleuGluIleGlnAlaGlyAsnAsnVal 236
298 TTATATACCATTTCCGGTGGTATATACACCGCTACTCTCGTGCAGCT 347
237 .....SerProPhePheGlyThrValGlnleuAsnTrpAsnPhsIstH 251
348 GAGCAATGTCGCGAC.....GTGATGCCACAGCTTAG 379
251 rHIsArgLeuValAlaAsnPheGluGluLeuTrpLysIleThrGlyIleA 268
380 CGATCATGTTGTTGTATGTAGGACGACGACATCTCTCCGAGCTGCAC 429
268 lAlleuLeuPheArgCys...SerSerProSerAsnAsnAlaAlaIle 283
430 TATTGCCCGCTGTATACGACCCGCTTGGAATAATAGCGCGCGCTGCA 479
284 ArgMetProleuaspelunIagly.....GluAsnAsnLysTYrAsnAs 298
480 CGATGTTACTGCACCT 495
298 pGlyGluThrCysThr 303

seq_name: swIsaprot_40:RIP2_BRydi
seq_documentation_block:
ID RIP2_BRydi STANDARD: PRT: 282 AA.
AC P98184: Q98J80:
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Ribosome-inactivating protein bryodin II precursor (rRNA N-
DE glycosidase) (EC 3.2.2.22) (BD2).
OS Bryonia dioica (red bryony)
OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae:
OC eurosoids I: Cucurbitales: Cucurbitaceae: Bryonia.
OX NCBI_Taxid=3652;
RN [1]
RP SEQUENCE FROM N.A.
RA Siegall C.B., Gawlak S.L., Marquardt H.;
RT Bryonia dioica.
RL Patent number US5597569, 28-JAN-1997.
RN [2]
RP SEQUENCE OF 22-42.
RC TISSUE=Root;
```

MEDLINE=95151812; PubMed=7849072;
RA Siegall C.B., Gawlak S.L., Chace D., Wolff E.A., Mixan B.,
RA Marquardt H.;
RT "Characterization of ribosome-inactivating proteins isolated from
RT Bryonia dioica and their utility as carcinoma-reactive
RT immunoconjugates."
RL Bloconj. Chem. 5:423-429(1994).
CC - FUNCTION: RIBOSOME-INACTIVATING PROTEIN OF TYPE 1, INHIBITS
CC PROTEIN SYNTHESIS IN ANIMAL CELLS (BY SIMILARITY).
CC - CATALYTIC ACTIVITY: Endonucleolytic cleavage of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC - SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY. TYPE 1
CC RIP SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

```

DR EMBL: I14238; ? NOT_ANNOTATED_CDS.
DR HSSP: P09989; 1MRJ.
DR InterPro: IPR001574; RIP.
DR Pfam: PF00161; RIP; 1.
DR PRINTS: PR00396; SHIGARICIN.
DR PROSITE: PS00275; SHIGA_RICIN; 1.
KW protein synthesis inhibitor; Hydrolase; Toxin; Multigene family;
KM Glycoprotein; Signal
FT SIGNAL 1
FT CHAIN 22 282 RIBOSOME-INACTIVATING PROTEIN BRYODIN II.
FT ACT_SITE 183 183 BY SIMILARITY.
FT CARBOHYD 25 25 N-LINKED (GLCNAC... ) (POTENTIAL).
FT SEQUENCE 282 AA; 30754 MW; C52BE2F6A873769C CRC64;
```

alignment_scores:
Quality: 166.50 Length: 150
Ratio: 1.716 Gaps: 9
Percent Similarity: 64.667 Percent Identity: 36.000

alignment block:
US-09-627-165b-15 x RIP2_BRydi ..

Align seg 1/1 to: RIP2_BRydi from: 1 to: 282

```

1 TACACAGATCTGAGGATACGCCGCTCAT.....AGGACCAAGATCCC 44
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
132 TyrapGlyleugluThralaAlaGlyArgIleSerArgGluAsnIleGl 148
45 TCGGGTATAGAGAACTCATTCATCCGCTCGCGCGCT.....CGTT 88
148 uLeuglyPheSerGluIleSerSerAlaIleGlyAsnMetPheArgHisA 165
89 ATCCAGCGCGGACGACCCGCGCCCAAGCTGCTTCCTTAATCTTCATTT 138
165 snProGlyThrSerVal.....ProArgAlaPheIleValIleIle 178
139 CAGATGATCTCGAGCGCGGAGATTCATCCATCTTTGGAGGGCTCG 188
179 GlnThrValSerGluAlaAlaArgPheLysTYrIleGluGluValIle 195
189 CCAATATATTAACAGCGGAGATTCCTCCGACATGTACATGCTCGC 238
195 rGlu.....AsnValGlyThrLysPheLysProAlaPheLeu 210
239 AGCTGAGAGTATTGGGCGCCAAATCCACGAGATCCAG.....CAG 282
210 erleuGlnAsnAlaIatrpGlySerleuSerGluInIleGlnIleAlaGln 226
283 TCATCGATGCGCGTTTATATACCATTTCCGGTGGGTATA.....TC 326
  ::| ||| ||| ::| ||| ::| |||
```

ID	seq_documentation_block	STANDARD	PRT	289 AA.
AC	P09989			
DT	01-MAR-1989 (Rel. 10, Created)			
DT	01-NOV-1990 (Rel. 16, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	Ribosome-inactivating protein alpha-trichosanthin precursor			
OS	(rRNA N-91ycosidase) (EC 3.2.2.22) (Alpha-rts).			
OS	Trichosanthes kirilowii (Mongolian snake-gourd).			
OC	Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta;			
OC	Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae;			
OC	eurosid 1; Cucurbitales; Cucurbitaceae; Trichosanthes.			
OX	NCBI_TaxId=3677;			
ON	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=MAXIMOWICZ;			
RC	MEDLINE=91153657; PubMed=1999291;			
RA	Shaw P.C., Yung M.H., Zhu R.H., Ho W.K.K., Ng T.B., Yeung H.W.;			
RT	"Cloning of trichosanthin cDNA and its expression in <i>Escherichia</i>			
RT	<i>coli</i> .";			
RL	gene 97:267-272(1991).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=MAXIMOWICZ; TISSUE=leaf;			
RC	MEDLINE=90256790; PubMed=2341400;			
RA	Chow T., Feldman R.A., Lovett M., Platak M.;			
RT	"Isolation and DNA sequence of a gene encoding alpha-trichosanthin, a			
RT	type I ribosome-inactivating protein.";			
RN	J. Biol. Chem. 265:8670-8674(1990).			
RN	[3]			
RP	SEQUENCE OF 24-270.			
RC	STRAIN=MAXIMOWICZ; TISSUE=tuberous root;			
RX	MEDLINE=90256789; PubMed=2341399;			
RA	Collins E.J., Robertus J.D., Lopresti M., Stone K.L., Williams K.R.,			
RA	Wu P., Hwang K., Platak M.;			
RT	"Primary amino acid sequence of alpha-trichosanthin and molecular			
RT	models for abrin A-chain and alpha-trichosanthin.";			
RL	J. Biol. Chem. 265:8665-8669(1990).			
RN	[4]			
RP	SEQUENCE OF 24-270.			
RC	TISSUE=tuberous root;			
RA	Wang Y., Qian R.Q., Gu Z.W., Jin S.W., Zhang L.Q., Xia Z.X.,			
RA	Tian G.Y., Ni C.Z.;			
RT	"Scientific evaluation of Tian Hua Fen (THF): history, chemistry and			
RT	application.";			
RL	Pure Appl. Chem. 58:789-798(1986).			
RN	[5]			
RP	X-RAY CRYSTALLOGRAPHY (1.88 ANGSTROMS).			
RX	MEDLINE=94344957; PubMed=806085;			
RA	Zhou F., Fu Z., Chen M., Lin Y., Pan K.;			
RT	"Structure of trichosanthin at 1.88-A resolution.";			
RL	Proteins 19:4-13(1994).			
RN	[6]			
RP	X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).			
RX	MEDLINE=95344383; PubMed=7619070;			
RA	Huang Q., Liu S., Tang Y., Jin S., Wang Y.;			
RT	"Studies on crystal structures, active-center geometry and			
RT	deurinating mechanism of two ribosome-inactivating proteins.";			
RL	Biochem. J. 309:285-298(1995).			
CC	-1- FUNCTION: TRICHOSANTHIN IS AN ABORTION-INDUCING PROTEIN. IT IS			
CC	CAPABLE OF INHIBITING HIV-1 INFECTION AND REPLICATION. IT			

```

1  TACACAGATCTGGAGCCGATACGCCGGTCAT...AGGACACGATCCCTCT 47
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
134 TyrgtuatrgleugIntHrAlaAlaGlyLysIleargtGusmIleProle 150
48 GGGRTAGAGGAACCTCATTCGCTCGCGCCTTGATATCCAGGCG 97
   ||||| ||| ||| ::::: ||||| ||| ||| ||| :::::
150 uGlyLeuproAlaLeuaspSerAlaIlethrThrLeupheryIlyrYrsmA 166
98 GCAGACGCCGGGCCCAAGCTCGTCCCTTATATCCATTCACAGATATC 147
   ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
167 laasSerAlaAlaSerAla....LeuMetValIleuIleGlnSerThr 181
148 TCCGAGGCCCGGAGATTCATCCCATCTTTTGGAGGGCTGCCAATATCAT 197
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
182 SerLuuAlaAlaIargYrLysPheIleGlnGlnGlnIleGlyLysArgVa 198

```


RC	TISSUE=Seed:
RX	MEDLINE-91242096; PubMed-1368643;
RA	Kung S.S., Kimura M., Funatsu G.;
RT	"The complete amino acid sequence of antiviral protein from the seeds
RI	of pokeweed (<i>Phytolacca americana</i>).";
RL	Agric. Biol. Chem. 54:3301-3318.(1990).
CC	-1- FUNCTION: INHIBITS VIRAL INFECTIION OF PLANTS.
CC	SANTHERS IN VITRO.
CC	-1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at on
CC	specific adenosine on the 28S rRNA.
CC	-1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY, TYPE 1
CC	RIP SUPERFAMILY.
DR	P1R; J05401; J05401.
DR	HSSP; Q03464; IAPA.
DR	InterPro; IP001574; RIP.
DR	Pfam; PF00161; RIP; 1.
DR	PRINTS; PR00396; SHIGARGICIN.
DR	PROSITE; PS00275; SHIGA_RICIN; 1.
KW	Antiviral; Protein synthesis inhibitor; Hydrolase; Toxin.
FT	ACT_SITE 175 175
FT	DISULFID 34 258
FT	DISULFID 84 105
FO	SEQUENCE 261 AA; 29200 MW; D88B9962PEE8399D CRC64.

```
Align seg 1/1 to: RIPS_PHYAM from: 1 to: 261
```

```

220 GlyAlaLeuProLysProLeuGluLeuLysAsnAlaAspGlyThrIysTr 236
342 GACCTTGAGCAATGTTCCGACGCGTATGCCACCTTAGCAGTCATTGTGT 391
    ::: :||| ::::: :::::::::::::::::::: :
236 PileValLeuArgValAspGluIleLysProAspValGlyLeuLeuAsnT 253
392 TTGTA 396
    :|||
253 yval 254

seq_name: SwissProt_40:RIPA_PHYAM

seq_documentation_block:
ID RIPA_PHYAM STANDARD; PRT; 294 AA.
AC 003464:
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Antiviral protein alpha precursor (PAP-alpha) (Ribosome-inactivating
protein) (RNA N-glycosidase) (EC 3.2.2.22).
OS Phytolacca americana (Common pokeweed) (Virginian pokeweed).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Caryophyllales; Caryophyllales; Phytolaccaceae; Phytolacca.
OX NCBI_TaxId=3527;
[1]
SEQUENCE FROM N.A.
RC TISSUE=Seed, leaf, and Root;
RX MEDLINE=93099240; PubMed=1281438;
RA Kataoka J., Habuka N., Masuta C., Miyano M., Koiwai A.;
RT "Isolation and analysis of a genomic clone encoding a pokeweed
antiviral protein.";
RL Plant Mol. Biol. 20:879-886(1992).
[2]
X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
MEDLINE=95010127; PubMed=7925458;
AGO H., Kataoka J., Tsuge H., Hakuba N., Inagaki E., Noma M.,
Miyano M.;
RT "X-ray structure of a pokeweed antiviral protein, coded by a new
genomic clone, at 0.23-nm resolution. A model structure provides a
suitable electrostatic field for substrate binding.";
RL Eur. J. Biochem. 225:369-374(1994).
-1- FUNCTION: INHIBITS VIRAL INFECTION OF PLANTS, AND PROTEIN
SYNTHESIS IN VITRO. HAS ALSO BEEN SHOWN TO INHIBIT THE
REPLICATION OF MAMMALIAN VIRUSES. THE PROTEIN MAY PROVIDE A
MEANS OF CELLULAR SUICIDE UPON INVASION BY A VIRUS.
-1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
specific adenosine on the 28S rRNA.
-1- SUBUNIT: MONOMER.
-1- SUBCELLULAR LOCATION: Cell wall.
-1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY. TYPE 1
RIP SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; D10600; BAA01451.1; -.
DR PIR; S28421; S28421.
DR PDB; IAPA; 3I-JAN-94.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00161; RIP; 1.
DR ProSITE; PS00275; SHIGA_RICIN; 1.
KW Antiviral; Protein synthesis inhibitor; Hydrolyase; Toxin; Signal;
KW Cell wall; 3D-structure.
FT SIGNAL 1..24
FT CHAIN 25..285 BY SIMILARITY.
FT PROPEP 286..294 ANTIVIRAL PROTEIN ALPHA.

```

```

FT DISULFID 58 282
FT DISULFID 108 130
FT STRAND 28 31
FT HELIX 32 34
FT HELIX 37 51
FT STRAND 57 58
FT STRAND 59 60
FT STRAND 61 63
FT TURN 67 68
FT STRAND 73 79
FT TURN 81 82
FT STRAND 85 91
FT TURN 92 94
FT STRAND 97 104
FT TURN 105 106
FT STRAND 107 113
FT TURN 114 115
FT HELIX 119 129
FT STRAND 136 138
FT STRAND 141 141
FT HELIX 147 154
FT TURN 155 155
FT HELIX 158 160
FT STRAND 163 163
FT HELIX 165 175
FT TURN 176 177
FT HELIX 183 195
FT TURN 196 197
FT HELIX 198 202
FT HELIX 204 212
FT TURN 213 215
FT STRAND 218 218
FT HELIX 222 240
FT STRAND 242 242
FT TURN 243 244
FT STRAND 245 253
FT TURN 255 256
FT STRAND 259 264
FT HELIX 265 271
FT STRAND 275 275
SQ SEQUENCE 294 AA: 33069 MW: F2EC27724F85596 CRC64:

```

```

alignment_scores:
  Quality: 156.00      Length: 134
  Ratio: 1.660         Gaps: 3
  Percent Similarity: 70.149   Percent Identity: 30.597

```

alignment_block:
US-09-627-165b-15 x RIPA_PHYAM ..

Align seg 1/1 to: RIPA_PHYAM from: 1 to: 294

```

1 TACACAGATCTGGAGCAGTACGCCGCTCAT...AGGACACAGATCCCTCT 47
   ||| |||||: :|||: :|||: :|||: ||
147 TyrProAlaLeuGluLysLysValGlyArgProArgSerGlnValGlnLe 163
48 GGGTATGAGGACTCATTCATCCGTCGCGCTTCGTTATCCAGGCG 97
   |||||: :||| :|||: :|||: :|||: :||
163 uGlyLeuGlnLeuAsnSerGlyIleGlyLysIleTyrGlyValAspS 180
98 GCACACACC...CGGCCCAAGCTCGTCCCTTAATCTCATTCAGAGT 144
   :||| :|||: :|||: :|||: :|||: :|||
180 erPheThrGluLysThrGluAlaGluPheLeuValAlaIleGlnMet 196
145 ATCTCCAGGCGCGAGATTCATCCATCCATCTTTGGAGGCGCAATA 194
   :|||: :|||: :|||: :|||: :|||: :|||
197 ValSerGluAlaAlaArgPheLysTyrIleGluAsnGlnValLys... 211
195 CATTAACAGCGGAGTATCTTCCCGACATGTACATGCTCGAGCTGG 244
   ||| :|||: :|||: :|||: :|||: :|||
212 .ThrAsnPheAsnArgAlaPheTyrProAsnAlaLysValLeuAsnLeuG 288

```

```

245 AGACTAGTGGGGCCCAATTCACGCACGACGATGCTACGATGCGC 294
   || |||||: :|||: :|||: :|||: :|||
228 IuGluSerTrpGlyLysIleSerThrAlaIleHisAsnAlaLysAsnGly 244
295 GTTTTATATACCATTCGTTGGGTATATATCCACCGGATACCTGTCAC 344
   :|||: :|||: :|||: :|||: :|||: :|||
245 AlaLeuThrSerProLeuGluLeuLysAsnAlaLysAsnGlySerIle 261
345 GTTGAGCATGTTGCGACGATCGCCAGCTTAGCATGATGTTGTTG 394
   :|||: :|||: :|||: :|||: :|||: :|||
261 eValLeuArgValAspAspIleGluProAspValGlyLeuLeuLysTyr 278
395 TA 396
||
278 al 278

seq_name: SwissProt_40:RIP1_MOMCH

seq_documentation_block:
ID RIP1_MOMCH STANDARD; PRT; 286 AA.
AC P16094; P24697;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Ribosome-inactivating protein momordin I precursor (rRNA
DE N-glycosidase) (EC 3.2.2.22) (Alpha-momorcharin) (Alpha-MMC).
OS Momordica charantia (Bitter melon) (Balsam pear).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucotsids I; Cucurbitales; Cucurbitaceae; Momordica.
OX NCBI_Taxid=3673;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RX MEDLINE=91159486; PubMed=2001404;
RA Ho W.K.K., Liu S.C., Shaw P.C., Yeung H.W., Ng T.B., Chan W.Y.;
RT "Cloning of the cDNA of alpha-momorcharin: a ribosome inactivating
RT protein."
RL Biochim. Biophys. Acta 1088:311-314(1991).
RN [2]
RP SEQUENCE OF 24-38.
RC TISSUE=Seed;
RX MEDLINE=89326691; PubMed=2753596;
RA Montecucchi P.C., Lazzarini A.M., Barbieri L., Stirpe F., Soria M.,
RA Lappi D.;
RT "N-terminal sequence of some ribosome-inactivating proteins."
RL Int. J. Pept. Protein Res. 33:263-267(1989).
RN [3]
RP SEQUENCE OF 24-70.
RC TISSUE=Seed;
RX MEDLINE=89005108; PubMed=3262509;
RA Casellas P., Dussosoy D., Falcasa A.I., Barbieri L., Stirpe F.,
RA Guillemot J.C., Ferrara P., Bolognesi A., Ceñal P., Stirpe F.;
RT "Trichosanthin, a ribosome-inactivating protein from the seeds of
RT Trichosanthes kirilowii Maximowicz. Purification, partial
RT characterization and use for preparation of immunotoxins."
RL Eur. J. Biochem. 176:581-588(1988).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
RX MEDLINE=94356447; PubMed=8075985;
RA Ren J., Wang Y., Dong Y., Stuart D.I.;
RT "The N-glycosidase mechanism of ribosome-inactivating proteins
RT implied by crystal structures of alpha-momorcharin."
RL Structure 2:7-16(1994).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.16 ANGSTROMS).
RX MEDLINE=94192822; PubMed=8143869;
RA Husain J., Tickle I.J., Wood S.P.;
RT "Crystal structure of momordin, a type I ribosome inactivating
RT protein from the seeds of Momordica charantia."
RL FEBS Lett. 342:154-158(1994).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).

```

DR	EMBL; X57682; CAA40869.1;	-.
DR	PIR; S14273; RLPUQG.	
DR	PIR; S16490; S16490.	
DR	PDB; 1AHA; 22-JUN-94.	
DR	PDB; 1AHB; 22-JUN-94.	
DR	PDB; 1AHC; 22-JUN-94.	
DR	PDB; 1MOM; 31-MAY-94.	
DR	PDB; 1MRG; 07-FEB-95.	
DR	PDB; 1MRH; 07-FEB-95.	
DR	PDB; 1MRI; 07-FEB-95.	
DR	GlycoSuiteDB; P16094; -. InterPro; IPRO01574; RIP.	
DR	Pfam; PF00161; RIP; 1.	
DR	PRINTS; PR00396; SHIGARICIN.	
DR	PROSITE; PS00275; SHIGA_RICIN_1.	
KV	protein synthesis inhibitor; Hydrolase; Toxin; Signal; Glycoprotein; 3D-structure.	
KW	SIGNAL.	
FT	CHAIN	1 23
FT	CHAIN	24 269
FT	PROPEP	270 286
FT	ACT_SITE	183 183
FT	CARBOHYD	250 250
FT	N-LINKED (GLCNAC. . .).	
FT	/FTID-CAR_000082.	
FT	MISSING IN MATURE PROTEIN.	
FT	RIBOSOME-INACTIVATING PROTEIN MOMORDIN I.	
FT	SEQUENCE	286 AA; 31532 MW; E1B013ABEC216CF CRC64;

```

1  TACACATCATCTGAGCGATACCGCGGTAT ... AAGGACCAAGATCCCT 47
134  TyGluArgLeuGlnIleIleAlaGlyLysProAlaGlyLysIleProIle 150
48  GGGATATGAGGAACACTATTCATCCGTCTCGCGGCTCTGATCCAGCG 97
150  eGlyIleProAlaIleAspSerAlaIleSerThrIleLeuHis .....T 165
98  GCACGACACCGCGGCCACAGGTCTGCTCCCTTATATCTCATTCAGATGATC 147
165  yAspSerThrIleAlaIleAlaGlyAlaIleLeuValIleIleGlnThr 181
148  TCGAGGCGCCGCGATTCATCCCATCTTTGGAGGCGTCCCAATCAT 197
182  AlGlnAlaIleAlaArgLysThrIleGlnIleGlnIleGlnIleArgAl 198
198  TTAACACGCGGAGATCTATTTCTTCGCGCATACATGGCTGAGGTGAGA 247
198  aTyArgAspGlu .....ValProSerLeuAlaThrIleSerLeuIle 213

```

-1- BIOTECHNOLOGY: Especially useful as

-1- BIO TECHNOLOGY: Especially useful as immunotoxin for

pharmacological applications as it has low toxicity in rats and mice but is potent once inside target cells.
-1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY. TYPE 1 RIP SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC EMBL, 124020; -; NOT_ANNOTATED_CDS.
CC PIR, S16491; S16491.
CC PDB, 1BRV; 04-MAR-98.
DR InterPro: IPR001574; RIP.
DR Pfam: PF00161; RIP; 1.
DR PRINTS: PR00396; SHIGARICIN.
DR PROSITE: PS00275; SHIGA_RICIN; 1.
KW Protein synthesis inhibitor; Hydrolase; Toxin; 3D-structure;
KW Multigene family; Glycoprotein; Signal.
FT SIGNAL 1 23
FT CHAIN 24 270 RIBOSOME-INACTIVATING PROTEIN BRYODIN I.
FT PROPEP 271 290 MISSING IN MATURE PROTEIN.
FT ACT_SITE 183 183 BY SIMILARITY.
FT ACT_SITE 212 212
FT CARBOHYD 214 214 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 250 250 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT MUTAGEN 212 212 E->K: REDUCES ACTIVITY 10-FOLD.
FT CONFLICT 61 65 RSSIS -> LRHXI (IN REF. 3).
SQ SEQUENCE 290 AA; 31788 MW; E966CD9C031A42DB CRC64;

alignment_scores:
Quality: 145.50 Length: 135
Ratio: 1.565 Gaps: 5
Percent Similarity: 68.889 Percent Identity: 31.111

alignment_block:

US-09-627-165b-15 x RIP1_BRYDI ..

Align seg 1/1 to: RIP1_BRYDI from: 1 to: 290

```

1 TACACAGATCTGAGCATACGCCGCTCAT..AGGACGACATCCCTC 47
  ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
134 TyrcLunrgleuGlnThrAlaAlaGlyLysIleArgLysAlaIleProle 150
48 GGGTATAGAGAACTCAATTCATCCGCTCGCGCGCTTCGTTATCCAGCG 97
  ||||| ||| ||| ||||| ||||| ||||| ||||| |||||
150 uGlyLeuProAlaLeuAlaSerAlaIleThrLeuTyrTyrThr 167
98 GCACGACCGCGGCCAGCTCGTCCCTTAAATCCATCAATCCATGATC 147
  ||||| ||||| ||||| ||||| ||||| ||||| |||||
167 lAsSerAlaAlaSerAla.....LeuLeuValLeuIleGlnSerThr 181
148 TCCGAGCCGCGAGATCAATCCATCTTTGGAGCGCTGCACATATCAT 197
  ||||| ||||| ||||| ||||| ||||| ||||| |||||
182 AlaeLueAlaArgTyrLysPheIleGlnGlnGlnIleGlyLysArgVa 198
198 TAACAGCGGGGATCATTTCTCCGACATGATACATCTCGAGCTGAGA 247
  ||||| ||||| ||||| ||||| ||||| ||||| |||||
198 lAsP.....LysThrPheLeuProSerLeuAlaThrIleSerLeuAla 213
248 CTAGTGGGGCCACATCCAGCAGCATGTCAG.....CAGTCAGGAT 291
  ||||| ||||| ||||| ||||| ||||| ||||| |||||
213 snAsnThrPheAlaLeuSerLysIleGlnIleAlaSerThrAsnAsn 229
292 GCGCTTTTAAATACCATTTGCTGGGTATATACACGCTAACTTCGT 341
  ||||| ||||| ||||| ||||| ||||| ||||| |||||
230 GlylnPheGlnUserProValLeuIleAlaGlyLysAlaGlnArgVa 246
342 GACGTTGACGAA.....GTTCCGCGACGTGATCCGACGTTACGATCA 385

```

```

||||| ||| ||| ||||| |||||
246 lSerIleThrAsnAlaSerAlaArgValValThrSerAsnIleAlaLeuL 263
366 TGTGT 390
  ||||
263 euleu 264

```

seq_name: SwissProt_40:RIP1_PHYAM

seq_documentation_block:

```

ID RIP1_PHYAM STANDARD; PRT; 313 AA.
AC P10297;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Antiviral protein I precursor (PAP-I) (PAP-C) (Ribosome-inactivating
DE protein) (tRNA N-glycosidase) (EC 3.2.2.22).
GN PAP1.
OS Phytolacca americana (Common pokeweed) (Virginian pokeweed).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Phytolaccaceae; Phytolacca.
OX NCBI_TaxID=3527;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-62.
RC TISSUE=Leaf;
RX MEDLINE=92003676; PubMed=1912488;
RA Lin Q., Chen Z.C., Antoniw J.F., White R.F.;
RT "Isolation and characterization of a cDNA clone encoding the
RL anti-viral protein from Phytolacca americana.";
RN [2]
RP SEQUENCE OF 23-65.
RX MEDLINE=69193489; PubMed=2930487;
RA Barbieri L., Bolognesi A., Centi P., Falasca A.I., Minghetti A.,
RA Garofano L., Guicciardi A., Lappi D., Miller S.P.;
RT "Ribosome-inactivating proteins from plant cells in culture.";
RL Biochem. J. 257:801-807(1989).
RN [3]
RP SEQUENCE OF 23-54.
RC TISSUE=Leaf;
RX MEDLINE=83290867; PubMed=6885760;
RA Houston L.L., Ramakrishnan S., Hermodson M.A.;
RT "Seasonal variations in different forms of pokeweed antiviral protein,
RL a potent inactivator of ribosomes.";
RN [4]
RP SEQUENCE OF 23-54.
RC TISSUE=Leaf;
RX MEDLINE=85023392; PubMed=6091760;
RA Bjorn M.J., Larrick J., Platak M., Wilson K.J.;
RT "Characterization of transalation inhibitors from Phytolacca
RL americana. amino-terminal sequence determination and antibody-
inhibitor conjugates.";
RN [5]
RP SEQUENCE OF 23-54.
RC TISSUE=Root;
RX MEDLINE=91064383; PubMed=2248976;
RA Bolognesi A., Barbieri L., Abdondanza A., Falasca A.I., Carnicelli D.,
RA Battelli M.G., Stirpe F.;
RT "Purification and properties of new ribosome-inactivating proteins
RL with RNA N-glycosidase activity.";
RN [6]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX MEDLINE=94016586; PubMed=8411176;
RA Monzingo A.F., Collins E.J., Ernst S.R., Irvin J.D., Robertus J.D.;
RT "The 2.5 A structure of pokeweed antiviral protein.";
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 23-284.
RX MEDLINE=99421320; PubMed=10493577;

```

RA Kurinov I.V., Myers D.E., Irvin J.D., Uckun F.M.;
 FT "X-ray crystallographic analysis of the structural basis for the
 FT interactions of pokeweed antiviral protein with its active site
 FT inhibitor and ribosomal RNA substrate analogs.";
 RL Protein Sci. 8:1765-1772(1999)
 CC -1- FUNCTION: INHIBITS VIRAL INFECTION OF PLANTS, AND PROTEIN
 CC SYNTHESIS IN VITRO.
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
 CC specific adenosine on the 28S rRNA.
 CC -1- TISSUE SPECIFICITY: SPRING LEAVES AND ROOTS.
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY, TYPE 1
 CC RIP SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X55383; CAA39054.1; -;
 DR PIR; S02792; S02792.
 DR PIR; S13469; S13469.
 DR PDB; 1PAF; 31-JAN-94.
 DR PDB; 1PAG; 31-JAN-94.
 DR PDB; 1QCI; 15-SEP-99.
 DR PDB; 1QCG; 15-SEP-99.
 DR PDB; 1QCF; 15-SEP-99.
 DR InterPro: IPR001574; RIP.
 DR Pfam: PF00161; RIP: 1.
 DR PRINTS: PR00396; SHICARICIN.
 DR PROSITE: PS00275; SHIGA_RICIN; 1.
 DR Activital; Protein synthesis inhibitor; Hydrolase; Toxin; Signal;
 KW 3D-structure.
 FT SIGNAL 1 22
 FT CHAIN 23 285
 FT PROPEP 286 313
 FT ACT_SITE 198 198
 FT DISULFID 56 281
 FT DISULFID 107 128
 FT STRAND 25 29
 FT HELIX 35 49
 FT STRAND 55 56
 FT STRAND 57 58
 FT STRAND 59 61
 FT STRAND 65 66
 FT STRAND 71 78
 FT HELIX 79 81
 FT STRAND 82 89
 FT STRAND 90 93
 FT STRAND 94 102
 FT STRAND 103 105
 FT STRAND 106 112
 FT TURN 113 114
 FT TURN 117 117
 FT HELIX 118 127
 FT TURN 131 132
 FT STRAND 134 136
 FT STRAND 139 139
 FT HELIX 145 152
 FT TURN 153 153
 FT TURN 157 158
 FT STRAND 162 162
 FT HELIX 164 174
 FT TURN 175 176
 FT HELIX 182 195
 FT TURN 196 196
 FT HELIX 197 201
 FT HELIX 203 211
 FT TURN 212 214
 FT STRAND 217 217
 FT HELIX 221 239

ANTIVIRAL PROTEIN I.
 BY SIMILARITY.

FT STRAND	241	241
FT TURN	242	243
FT STRAND	244	252
FT TURN	254	255
FT STRAND	258	263
FT HELIX	264	270
FT STRAND	274	274
SO SEQUENCE	313 AA;	35219 MW; 2C57B2861EBA57F5 CRC64;

alignment_scores:			
Quality:	144.50	Length:	135
Ratio:	1.606	Gaps:	3
Percent Similarity:	66.667	Percent Identity:	30.370

alignment block:
 US-09-627-165b-15 x RIP1_PHYM ..

Align seg 1/1 to: RIP1_PHYM from: 1 to: 313

```

1 TACACAGATCTGGAGCGATACGGCGGT.....CATAGGACCAAGATCCC 44
||| ||||| ||||| ||||| |||||
145 TyrProThrIeuGIuSerLysAlaGlyValLysSerArgSerGlnValGI 161
45 TCTGGTATAGAGGAACTCATTCATCCGTCGCGGCTTCGTTATCCAG 94
||||||| ||| :||| :||| :||| :||| :||| :||| :|||
161 nLeuGlyIleGlnIleLeuaspSerasnIleGlyLysIleSerGIyValM 178
95 GCGGCGACACC...CGGCGCCAGCTCGTCCCTTATATCTCATTCAG 141
:|: ||| :||| :||| :||| :||| :||| :||| :|||
178 eSerPheThrGluLysThrGluAlaGluPheLeuLeuValAlaIleGln 194
142 ATGATCTCCGAGCGCGAGATTCATCCATCCACTTTTGGAGGCGCTGGCA 191
|||:||||| ||||| ||||| ||||| ||||| :||| :||| :|||
195 MetValSerIuAlaIaIaArgPheLysTyrIleGluasnGlnValLys.. 210
192 ATACATTACAGCGGAGATCTCTCCGACATGATCATCTCGAC 241
||| :||| :||| :||| :||| :||| :||| :||| :|||
211 ....ThrAsnPheAsnArgIaIaPheAsnProAsnProLysValLeuAsnL 226
242 TCGAGACTAGTTGGGCGCAACATCCAGCAAGTCAGATCTACGAT 291
|||:||||| ||||| ||||| ||||| ||||| :||| :||| :|||
226 euGIuGIuThrPGLysIleSerThrAlaIleHisAspAlaLysAsn 242
292 GCGGTTTTATATACCATTCGTTGGGTGGATATCCACCGGTAACTTCT 341
|||||:||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
243 GIyValLeuProLysProLeuGIuLeuValAspAlaSerGIyAlaLysTr 259
342 GACGTTGAGCAATGTTCCGACGTGATCCAGCACTTACGATCATGTTGT 391
:|: :||| :||| :||| :||| :||| :||| :||| :|||
259 pIleValIleuArgValAspGIuIleLysProAspValAlaLeuLeuAsnT 276
392 TNGTA 396
276 yVal 277

```


Percent Similarity: 71.739 Percent Identity: 45.652

alignment_block:
US-09-627-165B-15 x Q9FV22 ..

Align seg 1/1 to: Q9FV22 from: 1 to: 549

```

1  TACACGATCTGTGAGCGATACGCCGT...CATAGGACCAATCCCTCT 47
   |||||||
115  TYRTHASPLEUGLGLYVALALAGLYGUAARGGLUGLULLEULE 131
   |||||||
48  GGGTATAGAGAACTCATTCATCCGCTCTCGCGCTTGTATCCAGC. 96
   |||||||
131  UGLYMETASPROLEUGLUSNALALLESERALEUTRPILLESERANL 148
   |||||||
97  ..GGCAGCACCCGGCCCAAGCTCGTTCCTTATATATCCATTCAGATG 144
   |||||||
148  EUNANGINGLHARGALALEUALARGSERLEULLEVALILEGLINMET 164
   |||||||
145  ATCTCGAGGCGCGGAGATTCACATCCATTTGGAGGCGTCCGCATA 194
   |||||||
165  VALALAGLUALAVALARGPHEARGPHEILEGLUTYRARGVALARGLYSE 181
   |||||||
195  CATTAACACGGGGAGTCATTTCTCCCGACATGATGCTCGAGCTGG 244
   |||||||
181  RILLESERARGALAGLUMETPHEARGPROASPROALAMELLEUSERLEUG 198
   |||||||
245  AGACTAGTTGGGGCCACAAATCCAGCAGTCAGATCGGAT... 291
   |||||||
198  LUSANLYSTRPSERLALEUSERASNALAVALINGLINSERANGLY 214
   |||||||
292  GCGGTTTTTAATACCATTTGGTGGTATATCCACCGGTAATCTGCT 341
   |||||||
215  GLYVALPHESESERPROVALIGLULEUARGSERLLESERANLYSPROVA 231
   |||||||
342  GACGTGACCAATGTTCCGAC...GTGATCGCAGCTTAGGATCAATGT 388
   |||||||
231  ITRYVALIGLYSERVALSERASPARGVALILLESERGLYLEUALALILEMETL 248
   |||||||
389  TGTTTGATGTAG...GACCGACCAATCTTCTCCAGCTGCGCTAT 432
   |||||||
248  EUPHELLECYASRGSETRHASPARGALASERSEASPGLPHELEASP 264
   |||||||
433  TGGCGCTGTGTCATACGACCCGCTCTGAAAT.....AG 467
   |||||||
265  HISMETLEUMETLEARGPROILEUVALASPVALALAGLUALALATH 281
   |||||||
468  CGGCGCGCTCGAGATGTACCTGACGCTCCGAAACCGCAGCGGCGCA 517
   |||||||
281  RAAPALASAPASNASPAPTRHCYSALASAPROGLUPROTHVALARGI 298
   |||
518  TC 519
   ||
298  LE 298

```

seq_name: sp_plant:Q94BWS

seq_documentation_block:

```

ID   Q94BWS;          PRELIMINARY;      PRT;       581 AA.
AC   Q94BWS;
DT   01-DEC-2001 (TREMBLREL. 19, Created)
DT   01-DEC-2001 (TREMBLREL. 19, Last sequence update)
DE   01-DEC-2001 (TREMBLREL. 19, Last annotation update)
DE   TYPE 2 RIBOSOME-INACTIVATING PROTEIN CINNAMOMIN I PRECURSOR.
OS   Cinnamomum camphora (Camphor tree).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.
OX   NCBI_TaxId=13429;
RN   [1]
RP   SEQUENCE FROM N. A.
RA   Yang Q., Gong Z.Z., Liu W.Y.;
RT   Molecular cloning of three type 2 RIP (ribosome-inactivating protein)
RT   genes encoding cinnamomin proteins and study of their expression

```

patterns.":
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY039801; AAK82458.1; ..

FT	SIGNAL	1	32	POTENTIAL.
FT	CHAIN	33	581	TYPE 2 RIBOSOME-INACTIVATING PROTEIN CINNAMOMIN I
FT	SEQUENCE	581 AA;	64215 MW;	6E8F5FB8FBA3D196 CRC64;

alignment_scores:

Quality:	309.50	Length:	184
Ratio:	2.345 <td>Gaps:</td> <td>6</td>	Gaps:	6
Percent similarity:	71.739	Percent identity:	45.652

alignment_block:
US-09-627-165B-15 x Q94BWS ..

Align seg 1/1 to: Q94BWS from: 1 to: 581

```

1  TACACGATCTGTGAGCGATACGCCGT...CATAGGACCAATCCCTCT 47
   |||||||
147  TYRTHASPLEUGLGLYVALALAGLYGUAARGGLUGLULLEULE 163
   |||||||
48  GGGTATAGAGAACTCATTCATCCGCTCTCGCGCTTGTATCCAGC. 96
   |||||||
163  UGLYMETASPROLEUGLUSNALALLESERALEUTRPILLESERANL 180
   |||||||
97  ..GGCAGCACCCGGCCCAAGCTCGTTCCTTATATATCCATTCAGATG 144
   |||||||
180  EUNANGINGLHARGALALEUALARGSERLEULLEVALILEGLINMET 196
   |||||||
145  ATCTCGAGGCGCGGAGATTCACATCCATTTGGAGGCGTCCGCATA 194
   |||||||
197  VALALAGLUALAVALARGPHEARGPHEILEGLUTYRARGVALARGLYSE 213
   |||||||
195  CATTAACACGGGGAGTCATTTCTCCCGACATGATGCTCGAGCTGG 244
   |||||||
213  RILLESERARGALAGLUMETPHEARGPROASPROALAMELLEUSERLEUG 230
   |||||||
245  AGACTAGTTGGGGCCACAAATCCAGCAGTCAGATCGGAT... 291
   |||||||
230  LUSANLYSTRPSERLALEUSERASNALAVALINGLINSERANGLY 246
   |||||||
292  GCGGTTTTTAATACCATTTGGTGGTATATCCACCGGTAATCTGCT 341
   |||||||
247  GLYVALPHESESERPROVALIGLULEUARGSERLLESERANLYSPROVA 263
   |||||||
342  GACGTGAGCAATGTGCGGAC...GTGATCGCAGCTTAGGATCAATGT 388
   |||||||
263  ITRYVALIGLYSERVALSERASPARGVALILLESERGLYLEUALALILEMETL 280
   |||||||
389  TGTTTGATGTAG...GACCGACCAATCTTCTCCAGCTGCGCTAT 432
   |||||||
280  EUPHELLECYASRGSETRHASPARGALASERSEASPGLPHELEASP 296
   |||||||
433  TGGCGCTGTGTCATACGACCCGCTCTGAAAT.....AG 467
   |||||||
297  HISMETLEUMETLEARGPROILEUVALASPVALALAGLUALALATH 313
   |||||||
468  CGGCGCGCTCGAGATGTACCTGACGCTCCGAAACCGCAGCGGCGCA 517
   |||||||
313  RAAPALASAPASNASPAPTRHCYSALASAPROGLUPROTHVALARGI 330
   |||
518  TC 519
   ||
330  LE 330

```

seq_name: sp_plant:Q41174

seq_documentation_block:

```

ID   Q41174;          PRELIMINARY;      PRT;       541 AA.
AC   Q41174;

```


[illegible][illegible]

```
RT "Cloning and expression of three abrin A-chains and their mutants
RT derived by site-specific mutagenesis in Escherichia coli."
RL Eur. J. Biochem. 219:83-87(1994).
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RRNA.
CC -1- SIMILARITY: TO RIBOSOME INACTIVATING PROTEIN FAMILY.
DR EMBL: X76721; CAA54139.1; -.
DR HSSP: P11140; IABR.
DR InterPro: IPR001574; RIP.
DR Pfam: PF00161; RIP; 1.
DR PRINTS: PR00396; SHIGARICIN.
KW Hydrolase; Toxin.
FT NON_TER 1 1
FT NON_TER 251 251
SQ SEQUENCE 251 AA; 28040 MW; D57CB182E0EC9 CMC64;

alignment_scores:
  Quality: 243.00      Length: 145
  Ratio: 2.305        Gaps: 3
  Percent Similarity: 66.897      Percent Identity: 40.000

alignment_block:
US-09-627-165B-15 x Q96236 ..

Align seg 1/1 to: Q96236 from: 1 to: 251

1 TACACAGATCTGAGCGATACCGCGATCAT...AGGACGACAGATCCCTCT 47
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
113 TYRGLASPLEUGLARGTRPALHISGLINSEARGINGLILEPROLE 129
48 GGGTATAGAGGAGCTCATTCATCCGTCGCGCGCTGCTATACGAGCGG 97
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
129 UGLYLEUGLINALLEUTHRHISGLYILESERPHEPHARGSERGLYGLYA 146
98 GAGACGCCGGGCCAAGCTGCTCCCTTATATCCATTCAGATGAGATGC 147
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
146 SNAASPANGIUGLWYSALAARGTRHLEULLEVALILEGLINMETVAL 162
148 TCCGAGCGCGCGAGATTCATCCCATCTTTGGAGGGCTGCCAATACAT 197
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
163 ALAALALALALALARGPHEARGYRILESERASNAARGVALARGVALSER 179
198 TACAGCGGGGATCATTTCTCCGACATCTACATCTGAGCTGAGAGA 247
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
179 EGLINHRGLYTHRALPHEGLNPROASPALALAMETLILESERLEUGLUA 196
248 CTAGTTGGGGCCAAACATCCAGCAAGCTCCAGAGCTACGATGGCGTT 297
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
196 SNAASNTTPASPASNLEUSERARGLYVALGINGLINUSERVALGINASPTHR 212
298 TTATATTAACCATTTGCGTTGGGTATATCCACCGGTAACTTCGTGACGT 347
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
213 PHEPROASNGLN.....ValThrLe 219
348 GAGCAATGTTGCGGAC.....G 364
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
219 UTHRASNILIEARGASNGLUPROVALILEVALASPSERLEUSERHISPROT 236
365 TGATCGCCAGCTTAGCGATCATGTTGTTGTATGCT 399
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
236 hVALALVALLEUALALALEUKEULEUPHEVALGYS 247

seq_name: sp_plant:Q96237

seq_documentation_block:
ID Q96237 PRELIMINARY; PRT; 251 AA.
AC Q96237;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE tRNA - GLYCOSIDASE (EC 3.2.2.22) (FRAGMENT).
OS Abrus precatorius (Indian licorice) (Crab's eye).
```

```
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosid 1; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
OX NCBI_TaxID=3816;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SEED;
RA MEDLINE=94139756; PubMed=8307038;
RX Chih-Hung H., Lee M.C., Chen J.K., Lin J.Y.;
RT "Cloning and expression of three abrin A-chains and their mutants
RT derived by site-specific mutagenesis in Escherichia coli.";
RL Eur. J. Biochem. 219:83-87(1994).
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RRNA.
CC -1- SIMILARITY: TO RIBOSOME INACTIVATING PROTEIN FAMILY.
DR EMBL: X76722; CAA54140.1; -.
DR HSSP: P11140; IABR.
DR InterPro: IPR001574; RIP.
DR Pfam: PF00161; RIP; 1.
DR PRINTS: PR00396; SHIGARICIN.
KW Hydrolase; Toxin.
FT NON_TER 1 1
FT NON_TER 251 251
SQ SEQUENCE 251 AA; 28055 MW; 6F64755C3DEAFB79 CRC64;

alignment_scores:
  Quality: 242.00      Length: 145
  Ratio: 2.495        Gaps: 3
  Percent Similarity: 66.897      Percent Identity: 40.000

alignment_block:
US-09-627-165B-15 x Q96237 ..

Align seg 1/1 to: Q96237 from: 1 to: 251

1 TACACAGATCTGAGCGATACCGCGATCAT...AGGACGACAGATCCCTCT 47
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
113 TYRGLASPLEUGLARGTRPALHISGLINSEARGINGLILEPROLE 129
48 GGGTATAGAGGAGCTCATTCATCCGTCGCGCGCTGCTATACGAGCGG 97
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
129 UGLYLEUGLINALLEUTHRHISGLYILESERPHEPHARGSERGLYGLYA 146
98 GACACGCCGGGCCAAGCTGCTCCCTTATATCCATTCAGATGATGC 147
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
146 SNAASPANGIUGLWYSALAARGTRHLEULLEVALILEGLINMETVAL 162
148 TCCGAGCGCGCGAGATTCATCCCATCTTTGGAGGGCTGCCAATACAT 197
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
163 ALAAGLUALALALEUPHEARGYRILESERASNAARGVALARGVALSER 179
198 TACAGCGGGGAGATCATTTCTCCGACATCTACATCTGAGCTGAGAGA 247
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
179 EGLINHRGLYTHRALPHEGLNPROASPALALAMETLILESERLEUGLUA 196
248 CTAGTTGGGGCCAAACATCCAGCAAGCTCCAGAGCTACGATGGCGTT 297
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
196 SNAASNTTPASPASNLEUSERARGLYVALGINGLINUSERVALGINASPTHR 212
298 TTATATTAACCATTTGCGTTGGGTATATCCACCGGTAACTTCGTGACGT 347
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
213 PHEPROASNGLN.....ValThrLe 219
348 GAGCAATGTTGCGGAC.....G 364
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
219 UTHRASNILIEARGASNGLUPROVALILEVALASPSERLEUSERHISPROT 236
365 TGATCGCCAGCTTAGCGATCATGTTGTTGTATGCT 399
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
236 hVALALVALLEUALALALEUKEULEUPHEVALGYS 247

seq_name: sp_plant:Q41358
```

```

seq_documentation_block:
ID 041358      PRELIMINARY;      PRT;      570 AA.
AC 041358:
DT 01-NOV-1996 (TRENBLER. 01, Created)
DT 01-NOV-1996 (TRENBLER. 01, last sequence update)
DT 01-DEC-2001 (TRENBLER. 19, last annotation update)
DE RNA - GLYCOSIDASE PRECURSOR (EC 3.2.2.22).
OS Sambucus nigra (European elder).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Adoxaceae; Sambucus.
OX NCBI_TaxID=4202;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BARK;
RX MEDLINE=96202926; PubMed=8611319;
RA Van Damme E.J., Barre A., Rouge P., Van Leuven F., Peumans W.J.;
RT "The Neucalpha-2,6)-Gal/GalNAc-binding lectin from elderberry
RT (Sambucus nigra) bark, a type-2 ribosome-inactivating protein with an
RT unusual specificity and structure.";
RL Eur. J. Biochem. 235:128-137(1996).
CC -1- CATALYTIC ACTIVITY: ENDOPHYLOSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RNA.
CC -1- SIMILARITY: TO RIBOSOME INACTIVATING PROTEIN FAMILY.
DR EMBL: U27122; AAC49158.1; -.
DR HSP: P02879; ZMAI.
DR InterPro: IPR000772; Ricin_B_lectin.
DR InterPro: IPR001574; RIP.
DR Pfam: PF00652; Ricin_B_lectin; 2.
DR Pfam: PF00161; RIP; 1.
DR PRINTS: PR00396; SHIGARICIN.
DR SMART: SM00458; RICIN.
DR PROSITE: PS0231; RICIN_B_LECTIN; 2.
DR PROSITE: PS00275; SHIGA_RICIN; 1.
KW Hydrolase; Signal; Toxin.
FT SIGNAL 1 28 POTENTIAL.
FT CHAIN 29 308 LECTIN_A CHAIN.
FT CHAIN 309 570 LECTIN_B CHAIN.
SQ SEQUENCE 570 AA; 63101 MW; A059E2A3E86B868 CRC64;

alignment_scores:
Quality: 237.00      Length: 186
Ratio: 1.911      Gaps: 7
Percent Similarity: 66.667      Percent Identity: 34.946

alignment_block:
US-09-627-165B-15 x 041358

Align seg 1/1 to: 041358 from: 1 to: 570

1 TACACAGATCTGGAGCATAGCGGTCAT..AGGACAGCATCCCTCT 47
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
145 TTTTSSerLeuGlnArgGlnValGlyPheGlyArgValTyrIleProLe 161
48 GGTATATAGAGAACTCATTCATCCGCTCTGGCGCTTCT.....T 88
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
161 uGlyProLysSerLeuAspGlnAlaIleSerSerLeuArgThrTyrTrpL 178
89 ATTCAGCGCGACACCGCGCGCCCAACCTGTTCCCTTAATTCCTCAAT 138
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
178 eutrrAlaGlyAspThrLysProLeuAlaArgGlyLeuLeuValValIle 194
139 CAGATGATCTCCGAGCGCGAGATCAATCCATCTTTGGAGGGCTCG 188
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
195 GlnMetValSerGlnAlaIlaArgPheArgTyrIleGluLeuArgIlaAr 211
169 CCAATTAATTAACGCGGGAGTCTTTCCCGACATGACATGCTCG 238
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
211 gThSerIleThrAspAlaSerGlnPheThrProAspLeuLeuMetLeu 228
239 AGCTGGAAGTACTGTGGGCGCAACATCCAGCAAGTCCAGAGTCT... 285

```

```

seq_name: sp_plant:Q96235

seq_documentation_block:
ID 096235      PRELIMINARY;      PRT;      251 AA.
AC 096235:
DT 01-FEB-1997 (TRENBLER. 02, Created)
DT 01-FEB-1997 (TRENBLER. 02, last sequence update)
DT 01-DEC-2001 (TRENBLER. 19, last annotation update)
DE RNA - GLYCOSIDASE (EC 3.2.2.22) (FRAGMENT).
OS Arabus precatorius (Indian licorice) (Crab's eye).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC eurosids II; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
OX NCBI_TaxID=3816;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SEED;
RX MEDLINE=94139756; PubMed=8307038;
RA Chih-Hung H., Lee M.C., Chen J.K., Lin J.Y.;
RT "Cloning and expression of three abrin A-chains and their mutants
RT derived by site-specific mutagenesis in Escherichia coli.";
RL Eur. J. Biochem. 219:83-87(1994).
CC -1- CATALYTIC ACTIVITY: ENDOPHYLOSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RNA.
CC -1- SIMILARITY: TO RIBOSOME INACTIVATING PROTEIN FAMILY.
DR EMBL: X76644; CA54092.1; -.
DR HSP: P1140; IABR.
DR InterPro: IPR001574; RIP.
DR Pfam: PF00161; RIP; 1.
DR PRINTS: PR00396; SHIGARICIN.
KW Hydrolase; Toxin.
FT NON_TER 1 251
FT NON_TER 251 1
SQ SEQUENCE 251 AA; 27997 MW; 3B60351839AEB7E CRC64;

alignment_scores:
Quality: 236.00      Length: 145
Ratio: 2.458      Gaps: 3
Percent Similarity: 66.207      Percent Identity: 39.310

alignment_block:
US-09-627-165B-15 x Q96235

Align seg 1/1 to: Q96235 from: 1 to: 251

```



```

Alignment_scores:      Length: 186
Quality: 231.00      Gaps: 7
Ratio: 1.863
Percent Similarity: 66.667      Percent Identity: 33.871

alignment_block:
US-09-627-165B-15 x 022415 ..

Align seg 1/1 to: 022415 from: 1 to: 570

1 TACACAGATCTGGAGCGATATACGGCGGTCAAT...AGGACACAGATCCCTCT 47
||| :|||:|||||:||||| :||| ||| |||||
145 TTTTLeSerLeuGluArgGlnValGlyPheIleValTyrlleProle 161
48 GGGTATAGAGAACTCAATTCACCTCGTCCGGCCTTCGT.....T 88
||||:|||||:|||||:|||||:|||||
161 uGlyProLySerLeuAlaGlnAlaIleSerSerLeuArgThrTyrlle 178
89 ATTCAGCGCGGACGACNCCGGGGCCCAAGCTGATCCCTTAATATCCGAT 138
|||:|||||:|||||:|||||:|||||
178 euserAlaGlyAspThrLyProLeuAlaArgGlyLeuLeuValIle 194
139 CACATGATCTCCGAGCCGCGAGATTCATTCCTCATCTTTGGAGCGCTCG 188
|||||:|||||:|||||:|||||:|||||
195 GlnMetValSerGlnAlaAlaArgPheArgTyrlleGluLeuArgIleAr 211
211 gThrSerIleThrAspAlaSerGluIleuPheThrProAspLeuLeuMetLeu 228
189 CCAATCAATTAACAGCGGGGAGTCATCTTCTCCGACATGATCAATGCTCG 238
|||:|||||:|||||:|||||:|||||
211 gThrSerIleThrAspAlaSerGluIleuPheThrProAspLeuLeuMetLeu 228
239 AGCTTGAGACTAGTTGGGGCCAAACAATTCACGCAAGTCCAGAGGCT... 285
1228 eMeGlnAlaAsnArgTrpSerSerMetSerIleGlnGlnAlaGln 244
286 ACCGATGCCGCTTTTAATTAACCATTCCTGGTGGATATACACCGGTAA 335
|||||:|||||:|||||:|||||:|||||
245 ProGlyGlyIlePheProGlyValValGlnLeuArgAspIleuArgAsnAs 261

```

seq_name: sp_plant:004071

AC 01-Jul-1997 (TEMBrel, 04, Created)
 BI 01-Jul-1997 (TEMBrel, 04, last sequence update)
 DI 01-Jul-1997 (TEMBrel, 19, last annotation update)
 DI 01-Dec-2001 (TEMBrel, 19, last annotation update)
 RNA - GLYCOSIDASE PRECURSOR (EC 3.2.2.22).
 DE Sambucus nigra (European elder).
 OS Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids II; Adoxaceae; Sambucus.
 OX NCBI_TaxID:4202;
 OX [1]

RX			SEQUENCE FROM RAB.
RP	MEDLINE-97236787;	PubMed-9075659;	
RA	Van Damme E.J.M., Barre A., Fougé P., Van Leuven F., Peumans W.J.;		
RT	"Isolation and molecular cloning of a novel type 2 ribosome-		
RT	inactivating protein with an inactive B chain from elderberry		
RT	(Sambucus nigra) bark."		
RL	J. Biol. Chem. 272:8353-8360(1997).		
CC	-1 CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE		
CC	SPECIFIC ADENOSINE ON THE 28S RNA.		
CC	-1 SIMILARITY: TO RIBOSOME INACTIVATING PROTEIN FAMILY.		
DR	EMBL; U58357; AAC49672.1; .		
DR	HSSP; P02879; ZAAI.		
DR	InterPro; IPR000772; Ricin_B_lectin.		
DR	InterPro; IPR001574; RIP.		
DR	Pfam; PF00652; Ricin_B_lectin; 2.		
DR	Pfam; PF00161; Rip; 1.		
DR	PRINTS; PR00396; SHIGARICIN.		
DR	SMART; SMO0458; RICIN; 2.		
DR	PROSITE; PS00231; RICIN_B_LECTIN; 2.		
DR	PROSITE; PS00275; SHIGA_RICIN; 1.		
KW	Hydrolase; Signal; Toxin.		
FT	SIGNAL	1	POTENTIAL.
FT	CHAIN	26	RIBOSOME INACTIVATING PROTEIN, A AND B
FT	CHAIN	26	CHAIN.
FT	CHAIN	283	RIBOSOME INACTIVATING PROTEIN, A CHAIN.
FT	CHAIN	565	RIBOSOME INACTIVATING PROTEIN, B CHAIN.
SO	SEQUENCE	565 AA; 62348 MD; CD744CC00593A26 CRC64;	

```
alignment_block:
  rnc-09-627-165B-15 x 004071
```

Align seg 1/1 to: 004071 from: 1 to: 565

```

DR HSP:PI1140; IABR.
DR InterPro: IPR000772; Ricin_B_lectin.
DR InterPro: IPR001574; R1P.
DR Pfam: PF00652; Ricin_B_lectin; 6.
DR Pfam: PF00161; R1P; 1.
DR PRINTS: PR00366; SHIRGAC1CN.
DR SMART: SM00458; R1CN; 2.
DR PROSITE: PS50231; R1CN_B_1; 2.
DR PROSITE: PS00733; SHIRG_R1CN; UNKNOWN_1.
DR HydroLase; toxin.
SQ SEQUENCE 347 AA; 61248 MW; 355A325C354A1BD CRC64;

alignment_scores:
    Quality: 222.00      Length: 165
    Ratio: 1.965         Gaps: 5
    Percent Similarity: 61.081      Percent Identity: 34.595

alignment block:
US-09-627-165B-15 x Q9ME69 ..

Align seg 1/1 to: Q9ME69 from: 1 to: 547

1 TACACAGATCTGGAGCGATACCGCGATCAT..AGGACCAATCCCTCT 47
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
132 TYRSPSPLEUHLUYSRTPLALHSELHSELRATGLATRLATRLSERLE 148
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
48 GGGTATGAGCAACTCATTCATCCGCTGTCGGGCGCTCTGTTATCCAGCG 97
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
148 UGLYLEUGLUALALEHATGALNGINLYILEYSPHELEATGSRGCLYALA 165
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
98 GCACACACCGCGCGCCCAAGTCGTCCTCCCTTAAATCCCTAATGATGATC 147
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
165 ERASPSAPGUGLUGLUEALHATGRLHLEUILLLELLEGLNMEVAL 183
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
148 TCCGAGCGCGCGAGATTCATCAATCCATCTTTTGGAGGCGCTGCCATPAT 197
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
182 ALGLUALLALATRPHEATGTYVALSERLYLEUVALLELLESLERLE 198
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
198 TAAACGCGGGGAGCTCATTTCTCCGACATGTACATGCTCGAGCTGGAGA 247
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
198 USERASTATGALALAHAPHEGLNPRPASPPOPSERMETLEUSERLEUGLA 215
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
248 CTAGTTGGGGCCCAACATCCACGACGAGCTCCAGCATGTACGAGTGGCGT 297
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
215 SNTHTTRPGIUPROLEUSERATGALVALGHNHSTHRYALGINSAPHR 233
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
298 TTTATATACCATTTGGTGGGTGATATACACGCGTAACCTGTGACGCT 347
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
232 PHEPROGLNASN.....VALTHRL 238
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
346 GAGCAATGTCGGGAC.....G 364
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
238 ULLEASNAVALRGNGIUALRGVALVALSERLEUSERHISPROS 255
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
365 TGATCGCCAGCTTACGCGATCATGTTGTTATGTATGAGGACGACCAACT 411
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
255 ERVALSERLALALUALALEUWELTEUPHEVALLYLSANPROLEUHSNLA 277
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
415 TCTCTCGACGCTGGCGTATATGGCCGCGTGTCTATAGACGCCGCTTGAAA 466
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
272 THRLINSER.....PROLEUENLEIARGSERVALVALGNGI 284
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
465 TAGGAGGCGCGTGCAGATGTACTGTGACGCGCTCGCAACCCAGCGTGC 514
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
284 NSER.....LYSLIECYSSERSEHISTYRGILUPROTHVALA 297
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
515 GCATC 519
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
297 rgile 298

```



```

seq_documentation_block:
; Sequence 16, Application US/09627165B
; GENERAL INFORMATION:
; APPLICANT: KIM, Jong-Bae
; TITLE OF INVENTION: CRUDE EXTRACT FROM VISCUM album coloratum, AND PROTEINS
; FILE REFERENCE: Korean Mistletoe Lectin
; CURRENT APPLICATION NUMBER: US/09/627,165B
; CURRENT FILING DATE: 2000-07-27
; NUMBER OF SEQ ID NOS: 16
; SEQ ID NO 16
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Viscum album coloratum
; FEATURE:
; NAME/KEY: misc_feature
US-09-627-165B-16

alignment_scores:
Quality: 408.50      Length: 122
Ratio: 4.045         Gaps: 1
Percent Similarity: 82.787      Percent Identity: 69.672

alignment_block:
US-09-627-165B-13 x US-09-627-165B-16 ..

Align seg 1/1 to: US-09-627-165B-16 from: 1 to: 174

1 GCCAGATTCAATCCCATGNTGTGAGCGCTTCGCCGCCAATTAACAGTGG 50
|||||
53 AIAAGPHEASNPROLLEUPHETPRGALARGGLNTRYLLEASNSEI 69
|||||
69 YGLNGLNSETHRGGLNGLNGLNSETHRASPGLYVALPHEASNAN 102
|||||
101 GTGCGCAATTCACCCAGTCCAGTCGACGATGGCATTTTAATACC 150
|||||
86 YGLNGLNSETHRGGLNGLNGLNSETHRASPGLYVALPHEASNAN 102
|||||
101 CAAATAGATTGCAGATTTCGCCGCTTACCTTGTGACGNTGACAAATG 200
|||||
103 PROPHEATGLEUGLYILESETHRGGLNGLNGLNSETHRASPGLYVALPHEASNAN 119
|||||
201 TGGCGACGATTCGACGATTCGCGGCTTGTGATGTCGAAATGAGTGGT 250
|||||
119 IARGSPVALILEALSERLEUALILEMETLEUPHEVALCYSGLYGLVA 136
|||||
251 GGCGAATTCCTCTCTGACCCACCTTCGCGCTGCTCTCAAGTCCGTC 300
|||||
136 TGPRTSERSETRASPVALARGTYRTPROLEUVALILEATGPROVAL 152
|||||
301 GTGCGAT.....GGCGCAACGATGTCACCTGACCTGATTTTCCGAAC 341
|||||
153 LEUGLUAASERGLYALALVALASPVALTHRCYSTHALASERGLUPR 169
|||||
342 CACCGTGGCATCGTA 357
|||||
169 OTHRYVALARGILEVAL 174

seq_name: /cgn2_6/prodata/2/paa/US093_COMB.pep:US-09-347-064-2

seq_documentation_block:
; Sequence 2, Application US/09347064
; GENERAL INFORMATION:
; APPLICANT: Eck, Jurgen
; APPLICANT: Schmidt, Arno
; APPLICANT: Zinke, Holger
; TITLE OF INVENTION: Recombinant Fusion Proteins Based on
; TITLE OF INVENTION: Ribosome-Inactivating Proteins of the mistletoe Viscum
; FILE REFERENCE: 09282-5

```

```

; CURRENT APPLICATION NUMBER: US/09/347,064
; CURRENT FILING DATE: 1999-07-02
; EARLIER APPLICATION NUMBER: PCT/EP98/00009
; EARLIER FILING DATE: 1998-01-02
; EARLIER APPLICATION NUMBER: EP 97 10 0012.0
; EARLIER FILING DATE: 1997-01-02
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Viscum album
US-09-347-064-2

alignment_scores:
Quality: 311.00      Length: 85
Ratio: 4.380         Gaps: 0
Percent Similarity: 83.529      Percent Identity: 71.765

alignment_block:
US-09-627-165B-13 x US-09-347-064-2 ..

Align seg 1/1 to: US-09-347-064-2 from: 1 to: 252

1 GCCAGATTCAATCCCATGNTGTGAGCGCTTCGCCGCCAATTAACAGTGG 50
|||||
168 AIAAGPHEASNPROLLEUPHETPRGALARGGLNTRYLLEASNSEI 184
|||||
51 GSAGTCNTCTCCACCAACATGTACATCTCGACCTGACGACGATTTGGG 100
|||||
184 YALASERHEUPROSPVALIYMETLEUGLLEUGLUTHRSETRPG 201
|||||
101 GTGCGCAATTCACCCAGTCCAGTCGACGATGGCATTTTAATACC 150
|||||
201 YGLNGLNSETHRGGLNGLNGLNSETHRASPGLYVALPHEASNAN 217
|||||
151 CAAATAGATTGCAGATTTCGCCGCTTACCTTGTGACGNTGACAAATG 200
|||||
218 PROLEATGLEUALIETPRPGGLYASNPHEVALTHREUPHASNVA 234
|||||
201 TGGCGACGATTCGACGATTCGCGGCTTGTGATGTCGAAATGAGTGGT 250
|||||
234 IARGSPVALILEALSERLEUALILEMETLEUPHEVALCYSGLYGLVA 251
|||||
251 GGCGA 255
|||||
251 TGPRT 252

seq_name: /cgn2_6/prodata/2/paa/US093_COMB.pep:US-09-347-064-8

seq_documentation_block:
; Sequence 8, Application US/09347064
; GENERAL INFORMATION:
; APPLICANT: Eck, Jurgen
; APPLICANT: Schmidt, Arno
; APPLICANT: Zinke, Holger
; TITLE OF INVENTION: Recombinant Fusion Proteins Based on
; TITLE OF INVENTION: Ribosome-Inactivating Proteins of the mistletoe Viscum
; FILE REFERENCE: 09282-5
; CURRENT APPLICATION NUMBER: US/09/347,064
; CURRENT FILING DATE: 1999-07-02
; EARLIER APPLICATION NUMBER: PCT/EP98/00009
; EARLIER FILING DATE: 1998-01-02
; EARLIER APPLICATION NUMBER: EP 97 10 0012.0
; EARLIER FILING DATE: 1997-01-02
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 8
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Viscum album

```

US-09-347-064-8

alignment_scores: length:
 quality: 311.00 85
 ratio: 4.380 gaps: 0
Percent Similarity: 83.529 Percent Identity: 71.765

alignment_block:

US-09-627-165B-13 x US-09-347-064-8 ..

Align seg 1/1 to: US-09-347-064-8 from: 1 to: 252

```
1 GCCAGATTCATCCATCCTGAGAGGCTTCGCCGCAATTAACAGTGG 50
|||||
167 AlaArgPheAsnProIleuTrpArgAlaArgGlnTyrIleAsnSerG1 183
51 GGAGTCNTCTCCACCAACATGATGATGCTGAGCTGGAGACGAGTGGG 100
| ||| |||
183 yAlaSerPheLeuProAspValTyrMetLeuGlnLeuGlnThrSerTrpG 200
101 GTGGACATTCACCCAGTCAGCTTGCGCATGATGTTGCGAATGCAATAC 150
|||||
200 lYngInGlnSerThrGlnValGlnHisSerThrAspGlyValPheAsnAsn 216
151 CAATAGATTCAGATTCGCCGCGTAACCTTGTGACGNTGACCAATGT 200
|||||
217 ProIleArgLeuAlaIleProProGlyAsnPheValThrLeuThrAsnVa 233
201 TCGGACGATGATTCACAGCTTGCGCATGATGTTGCGAATGCAATGCTG 250
|||||
233 lArgAspValIleAlaSerLeuAlaIleMetLeuPheValCysGlyGln 250
251 GGCCA 255
|||||
250 rgPro 251
```

seq_name: /cgn2_6/ptodata/2/paa/US093_COMB.pep:US-09-347-064-2

seq_documentation_block:

```
; Sequence 2, Application US/09347064A
; GENERAL INFORMATION:
; APPLICANT: Eck, Jurgen
; APPLICANT: Zinke, Holger
; TITLE OF INVENTION: Recombinant Fusion Proteins Based on
; TITLE OF INVENTION: Ribosome-Inactivating Proteins of the mistletoe Viscum
; FILE REFERENCE: 09282-5
; CURRENT APPLICATION NUMBER: US/09/347,064A
; CURRENT FILING DATE: 1999-07-02
; EARLIER APPLICATION NUMBER: PCT/EP98/00009
; EARLIER FILING DATE: 1998-01-02
; EARLIER APPLICATION NUMBER: EP 97 10 0012.0
; EARLIER FILING DATE: 1997-01-02
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Viscum album
US-09-347-064-2
```

alignment_scores: length:
 quality: 311.00 85
 ratio: 4.380 gaps: 0
Percent Similarity: 83.529 Percent Identity: 71.765

alignment_block:

US-09-627-165B-13 x US-09-347-064-2 ..

Align seg 1/1 to: US-09-347-064-2 from: 1 to: 252

```
1 GCCAGATTCATCCATCCTGAGAGGCTTCGCCGCAATTAACAGTGG 50
|||||
168 AlaArgPheAsnProIleuTrpArgAlaArgGlnTyrIleAsnSerG1 184
51 GGAGTCNTCTCCACCAACATGATGATGCTGAGCTGGAGACGAGTGGG 100
| ||| |||
184 yAlaSerPheLeuProAspValTyrMetLeuGlnLeuGlnThrSerTrpG 201
101 GTGGACATTCACCCAGTCAGCTTGCGCATGATGTTGCGAATGCAATAC 150
|||||
201 lYngInGlnSerThrGlnValGlnHisSerThrAspGlyValPheAsnAsn 217
151 CAATAGATTCAGATTCGCCGCGTAACCTTGTGACGNTGACCAATGT 200
|||||
218 ProIleArgLeuAlaIleProProGlyAsnPheValThrLeuThrAsnVa 234
201 TCGGACGATGATTCACAGCTTGCGCATGATGTTGCGAATGCAATGCTG 250
|||||
234 lArgAspValIleAlaSerLeuAlaIleMetLeuPheValCysGlyGln 251
251 GGCCA 255
|||||
251 rgPro 252
```

seq_name: /cgn2_6/ptodata/2/paa/US093_COMB.pep:US-09-347-064-8

seq_documentation_block:

```
; Sequence 8, Application US/09347064A
; GENERAL INFORMATION:
; APPLICANT: Eck, Jurgen
; APPLICANT: Zinke, Holger
; TITLE OF INVENTION: Recombinant Fusion Proteins Based on
; TITLE OF INVENTION: Ribosome-Inactivating Proteins of the mistletoe Viscum
; FILE REFERENCE: 09282-5
; CURRENT APPLICATION NUMBER: US/09/347,064A
; CURRENT FILING DATE: 1999-07-02
; EARLIER APPLICATION NUMBER: PCT/EP98/00009
; EARLIER FILING DATE: 1998-01-02
; EARLIER APPLICATION NUMBER: EP 97 10 0012.0
; EARLIER FILING DATE: 1997-01-02
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Viscum album
US-09-347-064-8
```

alignment_scores: length:
 quality: 311.00 85
 ratio: 4.380 gaps: 0
Percent Similarity: 83.529 Percent Identity: 71.765

alignment_block:

US-09-627-165B-13 x US-09-347-064-8 ..

Align seg 1/1 to: US-09-347-064-8 from: 1 to: 252

```
1 GCCAGATTCATCCATCCTGAGAGGCTTCGCCGCAATTAACAGTGG 50
|||||
167 AlaArgPheAsnProIleuTrpArgAlaArgGlnTyrIleAsnSerG1 183
51 GGAGTCNTCTCCACCAACATGATGATGCTGAGCTGGAGACGAGTGGG 100
| ||| |||
183 yAlaSerPheLeuProAspValTyrMetLeuGlnLeuGlnThrSerTrpG 200
101 GTGGACATTCACCCAGTCAGCTTGCGCATGATGTTGCGAATGCAATAC 150
|||||
200 lYngInGlnSerThrGlnValGlnHisSerThrAspGlyValPheAsnAsn 216
```

```
151 CAATAGATTCAGATTCCCGGTAACCTTGAGCAGTACGATG 200
|||||  |||  |||||  |||||  |||||  |||||  |||||  |||||
217 ProileargleuAlaIleProProGlyAsnPhenValThrIleuThrAsnVa 233
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
201 TCGGACGTGATTCGACGTGGGATGATGTTGTCGATGACGATGTC 250
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
233 IArgAspValIleAlaSerIleuAlaIleMetIleuPheValCysIleuAsp 250
251 GGGCA 255
|||||
250 rgpProSerSer 251

seq_name: /cgn2_6/prodata/2/paa/US096_COMB.pep:US-09-627-165b-2
seq_documentation_block:
; Sequence 2, Application US/09627165B
; GENERAL INFORMATION:
; APPLICANT: KIM, Jong-Bae
; TITLE OF INVENTION: CRUDE EXTRACT FROM VISCUM album coloratum, AND PROTEINS
; FILE REFERENCE: Korean Mistletoe Lectin
; CURRENT APPLICATION NUMBER: US/09/627,165B
; CURRENT FILING DATE: 2000-07-27
; NUMBER OF SEQ ID NOS: 16
; SEQ ID NO 2
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Viscum album coloratum
; FEATURE:
; NAME/KEY: misc_feature
US-09-627-165b-2

alignment_scores:
Quality: 303.00      Length: 88
Ratio: 4.151        Gaps: 0
Percent Similarity: 82.955      Percent Identity: 69.318

alignment_block:
US-09-627-165b-13 x US-09-627-165b-2  ..

Align seg 1/1 to: US-09-627-165b-2 from: 1 to: 254

1 GCCAGATTCAATCCCATGTCGTGGAGGCTTCGCGCAAAATTAACTGG 50
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
167 AlaArgPheAsnProIleuThrArgAlaArgGlnTyrIleSerSerGI 183
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
51 GGAGTCNTCTCCACCAACAATGTACATGCTCGAGCTGAGAGAGATGGG 100
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
183 YGlySerPheIleuProAspThrTyrIleuGlnIleuGlnThrSerTrp 200
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
101 GTGACAAATCCACCAAGTCCAGATGTCGACGATGATTTTAATAC 150
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
200 YGlnGlnIleSerThrGlnValGlnIleSerThrAspGlyValIlePheAsnVa 216
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
151 CAATAGATTCGACATTCGCGGTAACCTTGTCGACGNTGACGATG 200
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
217 ProIleArgIleuThrIleSerThrGlyValIlePheValThrIleuSerAsnVa 233
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
201 TCGGACGTGATTCGACGTGGGATGATGTTGTCGATGACGATGTC 250
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
233 IArgAspValIleAlaSerIleuAlaIleMetIleuPheValCysIleuAsp 250
251 GGGCAATTCCTCT 264
|||||  |||||
250 rgpProSerSer 254

seq_name: /cgn2_6/prodata/2/paa/US096_COMB.pep:US-09-627-165b-4
seq_documentation_block:
; Sequence 6, Application US/09627165B
; GENERAL INFORMATION:
; APPLICANT: KIM, Jong-Bae
; TITLE OF INVENTION: CRUDE EXTRACT FROM VISCUM album coloratum, AND PROTEINS
; FILE REFERENCE: Korean Mistletoe Lectin
; CURRENT APPLICATION NUMBER: US/09/627,165B
; CURRENT FILING DATE: 2000-07-27
; NUMBER OF SEQ ID NOS: 16
; SEQ ID NO 6
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Viscum album coloratum
; FEATURE:
; NAME/KEY: misc_feature
US-09-627-165b-6

alignment_scores:
Quality: 298.00      Length: 88
Ratio: 4.139        Gaps: 0
Percent Similarity: 81.818      Percent Identity: 68.182

alignment_block:
US-09-627-165b-13 x US-09-627-165b-4  ..

Align seg 1/1 to: US-09-627-165b-4 from: 1 to: 254

1 GCCAGATTCAATCCCATGTCGTGGAGGCTTCGCGCAAAATTAACTGG 50
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
167 AlaArgPheAsnProIleuThrArgAlaArgGlnTyrIleSerSerGI 183
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
51 GGAGTCNTCTCCACCAACAATGTACATGCTCGAGCTGAGAGAGATGGG 100
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
183 YGlySerPheIleuProAspThrTyrIleuGlnIleuGlnThrSerTrp 200
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
101 GTGACAAATCCACCAAGTCCAGATGTCGACGATGATTTTAATAC 150
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
200 YGlnGlnIleSerThrGlnValGlnIleSerThrAspGlyValIlePheAsnVa 216
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
151 CAATAGATTCGACATTCGCGGTAACCTTGTCGACGNTGACGATG 200
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
217 ProIleArgIleuThrIleSerThrGlyValIlePheValThrIleuSerAsnVa 233
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
201 TCGGACGTGATTCGACGTGGGATGATGTTGTCGATGACGATGTC 250
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
233 IArgAspValIleAlaSer***AlaIleMetIleuPheValCysIleuAsp 250
251 GGGCAATTCCTCT 264
|||||  |||||
250 rgpProSerSer 254

seq_name: /cgn2_6/prodata/2/paa/US096_COMB.pep:US-09-627-165b-6
```

alignment_scores: length: 87
 quality: 298.00
 ratio: 4.197
 percent similarity: 81.609 percent identity: 68.966

alignment block:
 US-09-627-165b-13 x US-09-627-165b-6 ..

Align seg 1/1 to: US-09-627-165b-6 from: 1 to: 256

```

1 GCCAGATTCATCCATCCTGTCGGAGCGCTCGCGGCAATTAACAGTGG 50
  |||||
169 AAlarGpheaSnprolleuTrpAlaArgInTyrlleaSnsergl 185
  |||||
185 yvalserTyrlleuProaspValTyrlmetleuGlutleuGlnalaSerTrpG 202
  |||||
101 GTCGACATCCACCCCAAGTCGACGACATGCTCAAGAGATGCTTTTAATAC 150
  |||||
202 lYgInGlnserThrglnValGlnInserThraSpelYalAlpheasnaSn 218
  |||||
151 CAATTAAGATTCAGATTTCCGCGGTAACTTTGTGACGNTGACGAAATG 200
  |||||
219 ProlleuArgleuGlylleSerThnglyAsnPhenValTrpleuSerAsnVa 235
  |||||
201 TCGCGACGTGATTCGACGCTTGGCGATCATGTTGTCGAATGCAAGTGTC 250
  |||||
235 lArGaspValllealaSerleuGlylleMetValAlpheValCyAlarGAspA 252
  |||||
251 GGCCATTCGCC 261
  |||||
252 rGserSerSer 255
  |||||

```

seq_name: /cgn2_6/ptodata/2/paa/US081_COMB.pep:US-08-101-329-6

seq_documentation_block:

Sequence 6, Application US/08101329

GENERAL INFORMATION:

APPLICANT: Wood, Mark S.

APPLICANT: Gould, Robert M.

APPLICANT: Kelleher, Peter J.

APPLICANT: Wallace, Thomas L.

TITLE OF INVENTION: SINGLE CHAIN-IMMUNOTOXIN

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PREVENTING SECONDARY

TITLE OF INVENTION: CATARACTS

NUMBER OF SEQUENCES: 35

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish and Richardson

STREET: 2200 Sand Hill Road, Suite 100

CITY: Menlo Park

STATE: California

COUNTRY: US

ZIP: 94025

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/101,329

FILING DATE: 03-AUG-1993

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Rae-Venter, Barbara

REGISTRATION NUMBER: 32,750

REFERENCE/DOCKET NUMBER: 05936/021001

TELEPHONE: (415) 854-5277

TELEFAX: (415) 854-0875

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 554 amino acids

TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-101-329-6

alignment_scores: length: 94
 quality: 140.50
 ratio: 2.266
 percent similarity: 65.957 percent identity: 35.106

alignment block:
 US-09-627-165b-13 x US-08-101-329-6 ..

Align seg 1/1 to: US-08-101-329-6 from: 1 to: 554

```

1 GCCAGATTCATCCATCCTGTCGGAGCGCTCGCGGCAATTAACAGTGG 50
  |||||
472 AAlarGpheaSnprolleuTrpAlaArgInTyrlleaSnsergl 472
  |||||
456 AAlarGpheaSnprolleuTrpAlaArgInTyrlleaSnsergl 472
  |||||
51 GGAGTCNTCCACCCCAAGTCGACGACATGCTCAAGAGATGCTTTTAATAC 100
  |||||
472 nArGaspValllealaSerleuGlylleMetValAlpheValCyAlarGAspA 489
  |||||
101 GTCGACATCCACCCCAAGTCGACGACATGCTCAAGAGATGCTTTTAATAC 150
  |||||
489 lYrGpheaSnprolleuTrpAlaArgInTyrlleaSnsergl 505
  |||||
151 CAATTAAGATTCAGATTTCCGCGGTAACTTTGTGACGNTGACGAAATG 200
  |||||
506 ProlleuArgleuGlylleSerThnglyAsnPhenValTrpleuSerAsnVa 522
  |||||
201 TCGCGACGTGATTCGACGCTTGGCGATCATGTTGTCGAATGCAAGTGTC 250
  |||||
522 lserlleuileProlleuilellealaSerleuGlylleMetValAlpheValCyAlarGAspA 539
  |||||
251 GGCCA.....TTCTCCTCTCTGACAC 273
  |||||
539 rGpheaSnprolleuTrpAlaArgInTyrlleaSnsergl 549
  |||||

```

seq_name: /cgn2_6/ptodata/2/paa/US07_COMB.pep:US-07-787-567A-1

seq_documentation_block:

Sequence 1, Application US/07787567A

GENERAL INFORMATION:

APPLICANT: Better, Marc D.

APPLICANT: Carroll, Steve F.

TITLE OF INVENTION: Materials Comprising and Methods of

TITLE OF INVENTION: Preparation and Use for Ribosome-Inactivating Proteins

TITLE OF INVENTION: Bicknell

NUMBER OF SEQUENCES: 36

CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &

STREET: Two First National Plaza, 20 South Clark

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60603

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/787,567A

FILING DATE: 19911104

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Gruber, Lewis S.

REGISTRATION NUMBER: 30,060

REFERENCE/DOCKET NUMBER: 27129/30545


```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/136,389
FILING DATE: 18-AUG-1998
APPLICATION NUMBER: 08/646,360
FILING DATE: 13-MAY-1996
APPLICATION NUMBER: PCT/US94/05348
FILING DATE: 12-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/054,691
FILING DATE: 12-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/988,430
FILING DATE: 09-DEC-1992
PRIOR APPLICATION DATA: US 07/901,707
FILING DATE: 19-JUN-1992
PRIOR APPLICATION DATA: US 07/787,567
FILING DATE: 04-NOV-1991
APPLICATION NUMBER: US 07/787,567
ATTORNEY/AGENT INFORMATION:
NAME: Monicholas, Janet M.
REGISTRATION NUMBER: 32,918
REFERENCE/DOCKET NUMBER: 200-70.P4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/707-8889
TELEFAX: 312/707-9155
TELEX: 650 388-1248
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 267 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-610-838-1

alignment_scores:
Quality: 138.00      Length: 87
Ratio: 2.379         Gaps: 0
Percent Similarity: 66.667      Percent Identity: 35.632

alignment_block:
US-09-627-1658-13 x US-09-610-838-1  ..

Align seg 1/1  to: US-09-610-838-1  from: 1  to: 267

1  GCCAGATTCATCCATCCTGTCGGAGGCTGCCCGGCAATTACAGCTGG  50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  :
179  Alaaargpheglintyrllelgluglglstumeclatgthrafgillearytyras  195
51  GGAGTCNTCCCAACCAATGATGCATGCCTCGAGTGGAGACGAGTGGG  100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  :
195  naargayseralaproaspproseralallerhrlleuglunansertirpg  212
101  GTGCAACATCCACCACCAAGTCACGACAGTCCCAAGAGTGGCATTTTAAATAC  150
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  :
212  lyaarglauserthnalallelglngluserasnglndlyalalphealaser  228
151  CAATTAGATGGAGATTTCGCCGGTAACTTGTGACGTGACATGAGCATGT  200
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  :
229  proleglnleuglnaargatgasaaglyserlyspheaserlalyraspyva  245
201  TCGGACAGCTGATCTCCAGCTTGGCGATCATGTTCTTCGATGAGAGTGTC  250
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  :
245  lserlleutlelprollerlelalelaleuMetValtyrArgcyahalarop  262
251  GGCCATTCTCC  261
|||||
262  roproSerSer  265
|||||

seq_name: /cgn2_6/ptodata/2/paa/JS096_COMB_pep.us-09-668-419-1
seq_documentation_block:

```

```

Sequence 1, Application US/09668419
GENERAL INFORMATION:
APPLICANT: VITETTA, ELLEN S.
APPLICANT: GHETIE, VICTOR F.
APPLICANT: SMALLSHAW, JOAN E.
APPLICANT: BALDINA, ROXANA G.
TITLE OF INVENTION: RICIN A CHAIN MUTANTS LACKING ENZYMATIC ACTIVITY AS VACCINES TO
TITLE OF INVENTION: AGAINST AEROSOLIZED RICIN
FILE REFERENCE: USPD:723US
CURRENT APPLICATION NUMBER: US/09/668,419
CURRENT FILING DATE: 2000-09-22
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 267
TYPE: PPT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-668-419-1

```

```
alignment_scores:      138.00      Length:      87
                       Quality:
                       Ratio:      2.379      Gaps:      0
Percent Similarity:    66.667      Percent Identity: 35.632
```

```
alignment_block:
ns-09-637-165B-13 x ns-09-668-419-1
```

Align seg 1/1 to: US-09-668-419-1 from: 1 to: 267

[illegible]


```

281 gileVal 283
seq_name: /sgn2_6/prodata/1/Pas/US09_NEW_COMB pep:us-09-601-667b-1
seq_documentation_block:
  Sequence 1, Application: US/09601667b
  GENERAL INFORMATION:
    APPLICANT: Morris, Peter
    APPLICANT: Stiefel, Thomas
    APPLICANT: Voelter, Wolfgang
    APPLICANT: Welters, Peter
    TITLE OF INVENTION: Recombinant Mistletoe Lectins
    FILE REFERENCE: 29841/35636
    CURRENT FILING DATE: 2000-10-06
    PRIOR APPLICATION NUMBER: PCT/EP99/00696
    PRIOR FILING DATE: 1999-02-03
    PRIOR APPLICATION NUMBER: D 199 04 210.8
    PRIOR FILING DATE: 1998-02-03
    NUMBER OF SEQ ID NOS: 41
    SEQ ID NO 1
    LENGTH: 333
    TYPE: PTI
    ORGANISM: Artificial Sequence
  FEATURE:
    OTHER INFORMATION: mistletoe lectin
  NAME/KEY: SITE
  LOCATION: 15
  OTHER INFORMATION: product="Xaa is Asp or Glu"
  FEATURE:
    OTHER INFORMATION: /label= Xaa1
  NAME/KEY: SITE
  LOCATION: 63
  OTHER INFORMATION: product="Xaa is Gly or Gln"
  FEATURE:
    OTHER INFORMATION: /label= Xaa2
  NAME/KEY: SITE
  LOCATION: 66
  OTHER INFORMATION: product="Xaa is Ile or Val"
  FEATURE:
    OTHER INFORMATION: /label= Xaa3
  NAME/KEY: SITE
  LOCATION: 75
  OTHER INFORMATION: product="Xaa is Leu or Ala"
  FEATURE:
    OTHER INFORMATION: /label= Xaa4
  NAME/KEY: SITE
  LOCATION: 107
  OTHER INFORMATION: product="Xaa is missing"
  FEATURE:
    OTHER INFORMATION: /label= Xaa5
  NAME/KEY: SITE
  LOCATION: 113
  OTHER INFORMATION: product="Xaa is Asn or Thr"
  FEATURE:
    OTHER INFORMATION: /label= Xaa6
  NAME/KEY: SITE
  LOCATION: 117
  OTHER INFORMATION: product="Xaa is Pro or Thr"
  FEATURE:
    OTHER INFORMATION: /label= Xaa7
  NAME/KEY: SITE
  LOCATION: 134
  OTHER INFORMATION: product="Xaa is Asp or Glu"
  FEATURE:
    OTHER INFORMATION: /label= Xaa8
  NAME/KEY: SITE
  LOCATION: 141
  OTHER INFORMATION: product="Xaa is Ser or Thr"
  FEATURE:
    OTHER INFORMATION: /label= Xaa9
  NAME/KEY: SITE
  LOCATION: 145
  OTHER INFORMATION: product="Xaa is Phe or Tyr"
  FEATURE:
    OTHER INFORMATION: /label= Xaa10
  NAME/KEY: SITE
  LOCATION: 152
  OTHER INFORMATION: product="Xaa is Thr or Ala"
  FEATURE:
    OTHER INFORMATION: /label= Xaa11
  NAME/KEY: SITE
  LOCATION: 177
  OTHER INFORMATION: product="Xaa is Ala or Tyr"
  FEATURE:
    OTHER INFORMATION: /label= Xaa12
  NAME/KEY: SITE
  LOCATION: 180
  OTHER INFORMATION: product="Xaa is Tyr or Asp"
  FEATURE:
    OTHER INFORMATION: /label= Xaa13
  NAME/KEY: SITE
  LOCATION: 185
  OTHER INFORMATION: product="Xaa is Ala or Glu"
  FEATURE:
    OTHER INFORMATION: /label= Xaa14
  NAME/KEY: SITE
  LOCATION: 191
  OTHER INFORMATION: product="Xaa is Val or Met"
  FEATURE:
    OTHER INFORMATION: /label= Xaa15
  NAME/KEY: SITE
  LOCATION: 210
  OTHER INFORMATION: product="Xaa is Ile or Phe"
  FEATURE:
    OTHER INFORMATION: /label= Xaa16
  NAME/KEY: SITE
  LOCATION: 224
  OTHER INFORMATION: product="Xaa is Pro or Ser"
  FEATURE:
    OTHER INFORMATION: /label= Xaa17
  NAME/KEY: SITE
  LOCATION: 225
  OTHER INFORMATION: product="Xaa is Pro or Thr"
  FEATURE:
    OTHER INFORMATION: /label= Xaa18
  NAME/KEY: SITE
  LOCATION: 232
  OTHER INFORMATION: product="Xaa is Thr or Ser"
  FEATURE:
    OTHER INFORMATION: /label= Xaa19
  NAME/KEY: SITE
  LOCATION: 236
  OTHER INFORMATION: product="Xaa is Asp or Ser"
  FEATURE:
    OTHER INFORMATION: /label= Xaa20
  NAME/KEY: SITE
  LOCATION: 287
  OTHER INFORMATION: product="Xaa is Asn or Ser"
  FEATURE:
    OTHER INFORMATION: /label= Xaa21
  NAME/KEY: SITE
  LOCATION: 290
  OTHER INFORMATION: product="Xaa is Cys or Arg"
  FEATURE:
    OTHER INFORMATION: /label= Xaa22
  NAME/KEY: SITE
  LOCATION: 325
  OTHER INFORMATION: product="Xaa is Gly or Asn"
  FEATURE:
    OTHER INFORMATION: /label= Xaa23
  NAME/KEY: SITE
  LOCATION: 364
  OTHER INFORMATION: product="Xaa is Gly or Asp"
  FEATURE:
    OTHER INFORMATION: /label= Xaa24

```

```

FEATURE:
NAME/KEY: SITE
LOCATION: 426
OTHER INFORMATION: product= "Xaa is Gly or Gln"
FEATURE:
NAME/KEY: SITE
LOCATION: 435
OTHER INFORMATION: product= "Xaa is Val or Asp"
FEATURE:
NAME/KEY: SITE
LOCATION: 439
OTHER INFORMATION: product= "Xaa is Gln or Lys"
FEATURE:
NAME/KEY: SITE
LOCATION: 442
OTHER INFORMATION: product= "Xaa is Gly or missing"
FEATURE:
NAME/KEY: SITE
LOCATION: 464
OTHER INFORMATION: product= "Xaa is Arg or Lys"
FEATURE:
NAME/KEY: SITE
LOCATION: 480
OTHER INFORMATION: product= "Xaa is Ala or Gly"
FEATURE:
NAME/KEY: SITE
LOCATION: 481
OTHER INFORMATION: product= "Xaa is Gly or Ala"
FEATURE:
NAME/KEY: SITE
LOCATION: 483
OTHER INFORMATION: product= "Xaa is Ser or Gly"
FEATURE:
NAME/KEY: SITE
LOCATION: 484
OTHER INFORMATION: product= "Xaa is Gly or Ser"
FEATURE:
NAME/KEY: SITE
LOCATION: 500
OTHER INFORMATION: product= "Xaa is Asn or Ser or Thr or Lys"

alignment_scores:
Quality: 373.00      Length: 119
Ratio: 3.885        Gaps: 1
Percent Similarity: 80.672      Percent Identity: 65.546

alignment block:
US-09-627-165B-13 x US-09-601-667B-1 ..
Align seq 1/1 to: US-09-601-667B-1 from: 1 to: 533
1 GCCAGATTCATCCATCCTGTCGAGCTTCCGCGCAATTAACAGTGG 50
|||||
168 AlaArgPheAsnProIleuLeuTyrParg**ArgGln**IleAsnSerGcl 184

```

```

51 GGAGTCNTCTCCACCAACATGATACATGCTTCGAGCTTGACACAGATTGGG 100
| ||| |||||
184 y***SerPheLeuProasp***TyrMetLeuGluLeuGluThrSerTrp 201
101 GTCCGACATTCACCCCAAGTCGACGATGCAAGATGCGATTTATATACC 150
|||||
201 IyGlnGlnSerThrGlnValGlnHisSerThrAspGlyValPheAsnAsn 217
151 CAATTAAGATGCGATTCCTCCGCGGTAACCTTGACAGTGAACATGAT 200
|||||
218 Pro***ArgLeuAlaIle*****GlyAsnPheValThrLeu***AsnVa 234
201 TCCGAGATGATCTCCAGCTTGCGATGATGTTGCGAATGACAGTGC 250
|||||
234 Iarg***ValIleAlaSerLeuAlaIleMetLeuPheValCysGlyGluA 251
251 GCGCATTCCTCTCTTCGACACACCTTCGCGCTGCTCCTAAGTCCGTC 300
|||||
251 rPProSerSerSerAspValArgTyrTrpProLeuValIleArgProVal 267
301 GTGATGCGGCGCAAGATGTCACCTGCACTNTTCCGAACCCAGCTGCG 350
||| |||||
268 Ile.....AlaAspAspValThrCysSerAlaSerGluProThrValAr 282
351 CATCGTA 357
|||||
282 gIleVal 284

seq_name: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:US-09-601-667B-40
seq documentation block:
: Sequence 40, Application US/09601667B
: GENERAL INFORMATION:
: APPLICANT: Morris, Peter
: APPLICANT: Stiefel, Thomas
: APPLICANT: Voelter, Wolfgang
: TITLE OF INVENTION: Recombinant Mistletoe Lectins
: FILE REFERENCE: 29841/36636
: CURRENT APPLICATION NUMBER: US/09/601,667B
: PRIOR FILING DATE: 2000-10-06
: PRIOR APPLICATION NUMBER: PCT/EP99/00696
: PRIOR APPLICATION NUMBER: D 198 04 210.8
: PRIOR FILING DATE: 1998-02-03
: NUMBER OF SEQ ID NOS: 41
: SEQ ID NO 40
: LENGTH: 534
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: mistletoe lectin
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (15)..(15)
: OTHER INFORMATION: Xaa is Asp or Glu
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (63)..(63)
: OTHER INFORMATION: Xaa is Gly or Gln
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (66)..(66)
: OTHER INFORMATION: Xaa is Ile or Val
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (75)..(75)
: OTHER INFORMATION: Xaa is Leu or Ala
: NAME/KEY: SITE
: LOCATION: (114)..(114)
: OTHER INFORMATION: Xaa is Asn or Thr

```

FEATURE:
NAME/KEY: SITE
LOCATION: (118)..(118)
OTHER INFORMATION: Xaa is Pro or Thr
FEATURE:
NAME/KEY: SITE
LOCATION: (135)..(135)
OTHER INFORMATION: Xaa is Asp or Glu
FEATURE:
NAME/KEY: SITE
LOCATION: (142)..(142)
OTHER INFORMATION: Xaa is Ser or Thr
FEATURE:
NAME/KEY: SITE
LOCATION: (146)..(146)
OTHER INFORMATION: Xaa is Phe or Tyr
FEATURE:
NAME/KEY: SITE
LOCATION: (153)..(153)
OTHER INFORMATION: Xaa is Thr or Ala
FEATURE:
NAME/KEY: SITE
LOCATION: (178)..(178)
OTHER INFORMATION: Xaa is Ala or Tyr
FEATURE:
NAME/KEY: SITE
LOCATION: (181)..(181)
OTHER INFORMATION: Xaa is Tyr or Asp
FEATURE:
NAME/KEY: SITE
LOCATION: (186)..(186)
OTHER INFORMATION: Xaa is Ala or Glu
FEATURE:
NAME/KEY: SITE
LOCATION: (192)..(192)
OTHER INFORMATION: Xaa is Val or Met
FEATURE:
NAME/KEY: SITE
LOCATION: (220)..(220)
OTHER INFORMATION: Xaa is Ile or Phe
FEATURE:
NAME/KEY: SITE
LOCATION: (225)..(225)
OTHER INFORMATION: Xaa is Pro or Ser
FEATURE:
NAME/KEY: SITE
LOCATION: (226)..(226)
OTHER INFORMATION: Xaa is Pro or Thr
FEATURE:
NAME/KEY: SITE
LOCATION: (233)..(233)
OTHER INFORMATION: Xaa is Thr or Ser
FEATURE:
NAME/KEY: SITE
LOCATION: (237)..(237)
OTHER INFORMATION: Xaa is Asp or Ser
FEATURE:
NAME/KEY: SITE
LOCATION: (288)..(288)
OTHER INFORMATION: Xaa is Asn or Ser
FEATURE:
NAME/KEY: SITE
LOCATION: (291)..(291)
OTHER INFORMATION: Xaa is Cys or Arg
FEATURE:
NAME/KEY: SITE
LOCATION: (326)..(326)
OTHER INFORMATION: Xaa is Gly or Asn
FEATURE:
NAME/KEY: SITE
LOCATION: (365)..(365)
OTHER INFORMATION: Xaa is Gly or Asp
FEATURE:

NAME/KEY: SITE
LOCATION: (427)..(427)
OTHER INFORMATION: Xaa is Gly or Gln
FEATURE:
NAME/KEY: SITE
LOCATION: (436)..(436)
OTHER INFORMATION: Xaa is Val or Asp
FEATURE:
NAME/KEY: SITE
LOCATION: (440)..(440)
OTHER INFORMATION: Xaa is Gln or Lys
FEATURE:
NAME/KEY: SITE
LOCATION: (443)..(443)
OTHER INFORMATION: Xaa is Gly or missing
FEATURE:
NAME/KEY: SITE
LOCATION: (444)..(444)
OTHER INFORMATION: Xaa is Arg or Lys
FEATURE:
NAME/KEY: SITE
LOCATION: (465)..(465)
OTHER INFORMATION: Xaa is Cys or Ser or Val
FEATURE:
NAME/KEY: SITE
LOCATION: (481)..(481)
OTHER INFORMATION: Xaa is Ala or Gly
FEATURE:
NAME/KEY: SITE
LOCATION: (482)..(482)
OTHER INFORMATION: Xaa is Gly or Ala
FEATURE:
NAME/KEY: SITE
LOCATION: (484)..(484)
OTHER INFORMATION: Xaa is Ser or Gly
FEATURE:
NAME/KEY: SITE
LOCATION: (485)..(485)
OTHER INFORMATION: Xaa is Gly or Ser
FEATURE:
NAME/KEY: SITE
LOCATION: (494)..(494)
OTHER INFORMATION: Xaa is Gly or Tyr
FEATURE:
NAME/KEY: SITE
LOCATION: (501)..(501)
OTHER INFORMATION: Xaa is Asn or Ser or Thr or Lys
FEATURE:
NAME/KEY: SITE
LOCATION: (502)..(502)
OTHER INFORMATION: Xaa is Ser or Gly
FEATURE:
NAME/KEY: SITE
LOCATION: (503)..(503)
OTHER INFORMATION: Xaa is Leu or Pro
FEATURE:
NAME/KEY: SITE
LOCATION: (504)..(504)
OTHER INFORMATION: Xaa is Ala or Met
FEATURE:
NAME/KEY: SITE
LOCATION: (505)..(505)
OTHER INFORMATION: Xaa is Met or Val
FEATURE:
NAME/KEY: SITE
LOCATION: (534)..(534)
OTHER INFORMATION: Xaa is Pro or Phe
US-09-601-667B-40

alignment_scores:
Quality: 373.00 Length: 119
Ratio: 3.885 Gaps: 1

Percent Similarity: 80.672 Percent Identity: 65.546

alignment_block:

US-09-627-165b-13 x US-09-601-667b-40 ..

Align seg 1/1 to: US-09-601-667b-40 from: 1 to: 534

```

1 GCCAGATTCAATCCCATCNTGTGGAGGCTTCCGCCGCAATTAACAGTGG 50
|||||
169 AlArGpHeAsnPrOllLeuTrPaRg***ArgGln**IleAsnSerGI 185
51 GGAGTCNTCTCCACCAACATGTCATGCTCGAGCTGGAGAGAGTGGG 100
|||||
185 Y**SerPheLeuProAsp**TyrMetLeuGlnLeuGlnThrSetrPg 202
101 GTCGACATCCACCCAGTCCAGCAGTCCAGAGATGGCATTTTAAATACC 150
|||||
202 YgInGlnSerThrGlnValGlnHisSerThrAspGlyValPheAsnAsn 218
151 CAAATAGATTGCAGATTTCGCCGGTAATTTGTGACGNTGAGCAATGT 200
|||||
219 Pro**ArgLeuAlaIle***GlyAsnPheValThrLeu**AsnVa 235
201 TCGCAGCTGATCTCAGCTTGGCAGATCATGTTGTCGAATGCAGTGTG 250
|||||
235 LArg**ValIleAlaSerLeuAlaIleMetLeuPheValCysGlyGlna 252
251 GGCCATTCTCCTCTCTGACCACCCCTTCCGCCGCTCTCAAGTCCGTC 300
|||||
252 rGProSerSerSerAspValAlaGlyTrPProLeuValIleArgProVal 268
301 GTGGATGGCGCCAGATGTCACCTGCAGCTTTTCCGACCCACCGTGGC 350
|||||
269 Ile.....AlaAspAspValThrCysSerAlaSerGlnurProthrValAr 283
351 CATCGTA 357
|||||
283 gIleVal 285

```

seq_name: /cgn2_6/prodata/1/paa/US09_NEW_COMB.pep:US-09-601-667b-5

seq_documentation_block:

```

; Sequence 5, Application US/09601667B
; GENERAL INFORMATION:
; APPLICANT: Morris, Peter
; APPLICANT: Stiefel, Thomas
; APPLICANT: Voelter, Wolfgang
; APPLICANT: Welters, Peter
; TITLE OF INVENTION: Recombinant Mistletoe Lectins
; FILE REFERENCE: 29841/36636
; CURRENT APPLICATION NUMBER: US/09/601,667B
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: PCT/EP99/00696
; PRIOR FILING DATE: 1999-02-03
; PRIOR APPLICATION NUMBER: D 198 04 210.8
; NUMBER OF SEQ ID NOS: 41
; SEQ ID NO 5
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Lectin A2
US-09-601-667b-5

```

alignment_scores: Quality: 327.00 Length: 88
Ratio: 4.247 Gaps: 0
Percent Similarity: 87.500 Percent Identity: 73.864

alignment_block:

US-09-627-165b-13 x US-09-601-667b-5 ..

Align seg 1/1 to: US-09-601-667b-5 from: 1 to: 256

```

1 GCCAGATTCAATCCCATCNTGTGGAGGCTTCCGCCGCAATTAACAGTGG 50
|||||
169 AlArGpHeAsnPrOllLeuTrPaRgTyrArgGlnAspIleAsnSerGI 185
51 GGAGTCNTCTCCACCAACATGTCATGCTCGAGCTGGAGAGAGTGGG 100
|||||
185 YGlnSerPheLeuProAspMetTyrMetLeuGlnLeuGlnThrSetrPg 202
101 GTCGACATCCACCCAGTCCAGCAGTCCAGAGATGGCATTTTAAATACC 150
|||||
202 YgInGlnSerThrGlnValGlnHisSerThrAspGlyValPheAsnAsn 218
151 CAAATAGATTGCAGATTTCGCCGGTAATTTGTGACGNTGAGCAATGT 200
|||||
219 ProPheArgLeuAlaIleSerThrGlnValAsnPheValThrLeuSerAsnVa 235
201 TCGCAGCTGATCTCAGCTTGGCAGATCATGTTGTCGAATGCAGTGTG 250
|||||
235 LArgSerValIleAlaSerLeuAlaIleMetLeuPheValCysGlyGlna 252
251 GGCCATTCTCCTCT 264
|||||
252 rGProSerSerSer 256

```

seq_name: /cgn2_6/prodata/1/paa/US09_NEW_COMB.pep:US-09-601-667b-38

seq_documentation_block:

```

; Sequence 38, Application US/09601667B
; GENERAL INFORMATION:
; APPLICANT: Morris, Peter
; APPLICANT: Stiefel, Thomas
; APPLICANT: Voelter, Wolfgang
; APPLICANT: Welters, Peter
; TITLE OF INVENTION: Recombinant Mistletoe Lectins
; FILE REFERENCE: 29841/36636
; CURRENT APPLICATION NUMBER: US/09/601,667B
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: PCT/EP99/00696
; PRIOR FILING DATE: 1999-02-03
; PRIOR APPLICATION NUMBER: D 198 04 210.8
; NUMBER OF SEQ ID NOS: 41
; SEQ ID NO 38
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: mistletoe lectin A2
US-09-601-667b-38

```

alignment_scores: Quality: 327.00 Length: 88
Ratio: 4.247 Gaps: 0
Percent Similarity: 87.500 Percent Identity: 73.864

alignment_block:

US-09-627-165b-13 x US-09-601-667b-38 ..

Align seg 1/1 to: US-09-601-667b-38 from: 1 to: 256

```

1 GCCAGATTCAATCCCATCNTGTGGAGGCTTCCGCCGCAATTAACAGTGG 50
|||||
169 AlArGpHeAsnPrOllLeuTrPaRgTyrArgGlnAspIleAsnSerGI 185
51 GGAGTCNTCTCCACCAACATGTCATGCTCGAGCTGGAGAGAGTGGG 100
|||||
185 YGlnSerPheLeuProAspMetTyrMetLeuGlnLeuGlnThrSetrPg 202
101 GTCGACATCCACCCAGTCCAGCAGTCCAGAGATGGCATTTTAAATACC 150

```

seq_name: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:US-09-601-667B-39

; Sequence 39, Application US/09601667B
; GENERAL INFORMATION:

```

? APPLICANT: Morris, Peter
? APPLICANT: Stiefel, Thomas
? APPLICANT: Voelker, Wolfgang
? APPLICANT: Wellers, Peter
? TITLE OF INVENTION: Recombinant Mistletoe Lectins
? FILE REFERENCE: 29841/36636
? CURRENT APPLICATION NUMBER: US/09/601,667B
? CURRENT FILING DATE: 2000-10-06
? PRIOR APPLICATION NUMBER: PCT/EP99/00696
? PRIOR FILING DATE: 1999-02-03
? PRIOR APPLICATION NUMBER: D 198 04 210.8
? PRIOR FILING DATE: 1998-02-03
? NUMBER OF SEQ ID NOS: 41
? SEQ ID NO 39
? LENGTH: 253
? TYPE: PRT
? ORGANISM: Artificial Sequence
? FEATURE:
?
? OTHER INFORMATION: Mistletoe lectin A1 (matched)
?
? US-09-601-667B-39

```

```
alignment_scores:
  Quality: 317.00
  Ratio: 4.342
  Percent Similarity: 82.955
  Length: 88
  Gaps: 0
  Percent Identity: 71.591
```

US-09-627-165B-13 x US-09-601-667B-39

Align seg 1/1 to: US-09-601-667B-39 from: 1 to: 253

1	GCAGATTCCAAATCCGACATCNGTGTGAGCGCTTCCGCCGCAAAATTTAACAGTGG	50
166	AlaArgPheAsnProIleLeuThrPheGlyValArgGlnTyrIleAsnSerG1	1832
51	GGAGTCNTTCACCAACAAATGATACATGGTCGACGTGGAGACGAGTTGGG	1000
182	ValAsnPheLeuProAspAlaTyrMetLeuGlnIleuGlnIuhSerTrpG	1990
101	GTCGACAAATCACCACCAAGTCACGACAGTCCAGGATGGCATTTTAAATACC	1500
199	IyeGlnGlnSerThrGlnValGlnHisSerThrAspGlyValPheAsnAsn	2150
151	CAAAATTAAGATTGCAGATTTTCCGCCGGTAACATTGTGACAGNAGCAATGT	2000
216	ProIleArgPheAlaIleProProGlyAsnPheValThrLeuThrAsnVa	2320
201	TTCGGAGGTATGTCAGGTCGGCATATGTTGTTCGAATGCAGTGGTC	2500
232	ArgAspValIleAlaSerLeuAlaIleMetLeuPheValCysGlyLeuA	2490
251	GGCCATTCCTCCCTCT	264

249 rgProSerSer 253

seq_name: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:US-09-601-667B-37

seq_documentation_block:

Sequence 37, Application US/09601667B

GENERAL INFORMATION:

APPLICANT: Morris, Peter

APPLICANT: Voelter, Wolfgang

APPLICANT: Welters, Peter

; TITLE OF INVENTION: Recombinant Mistletoe Lectins

FILE REFERENCE: 29841/36636

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

PRIOR APPLICATION NUMBER: DGT 47006 000000

PRIOR FILING DATE: 1999-02-03

PRIOR APPLICATION NUMBER: D 198 04 210 8

PRIOR FILING DATE: 1998-02-03

; NUMBER OF SEQ ID NOS: 41

; SEQ ID NO 37

LENGTH: 254

ORGANISM: Artibeus fonscolombei

FEATURE: artificial sequence

OTHER INFORMATION: mistletoe lectin A1

US-09-601-667B-37

alignment_scores:

Quality:	317.00	Length:	88
----------	--------	---------	----

Ratio:	4.342
Gaps:	0

Percent Similarity: 82.955 Percent Identity: 71.591

alignment_block:

US-09-627-165B-13 x US-09-601-667B-37

Align seg 1/1 to: US-09-601-667B-37 from: 1 to: 254

```

1  GCCAGATTTCATCCCATCTCMTGCGAGAGCGTTGGCCGGCCAAATTAAACAGTGG  50
167  AAtargheAnsProIleLeuThrphgIyrahgIlyrIIleAasSergI  163
51  GGAGCTCMTTCACCAACAACATGTACATGCTCGAGCTGAGAGCAGATTGGG  100
183  yAlaSer.PheIeuProAasPValTyrlMetLeuGluIeuGluThrSerTrpg  200
101  GTGCACATCATCACCACCAAGTCACGAGCTCCAGAGNTGCGATTTTAAATAC  150
200  IyngInGInSerThrGlnValGlnhISerThrAspGlyValAlpheasnasn  216
151  CAATTAAGATTGCAGATTTCGCGCGGTAACTTTGTGACGNTGACCAATGT  200
217  ProIleArgIeuAlaIleProProGlyasnPheValThrIeuThrAsnVa  233
201  TCCGACAGTGATCTCCAGGTTGGCAGTATGTTGTTCGANTGCAGAGTGC  250
233  IArgAspValIleIleAsIerLeuAlaIleMetIeuPheValIcySelyGlna  250
251  GGCCATCTTCCTCT  264
250  rgrProSerSerSer  254

```

seq_name: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:US-09-601-667B-2

seq_documentation_block:

; Sequence 2, Application US/09601667B

; GENERAL INFORMATION:

APPLICANT: MORRIS, Peter

APPLICANT: Voelter, Thomas

APPLICANT: **Welters, Peter**

TITLE OF INVENTION: Recombinant Mistletoe Lectins
FILE REFERENCE: 29841/3636
CURRENT APPLICATION NUMBER: US/09/601,667B
CURRENT FILING DATE: 2000-10-06
PRIORITY APPLICATION NUMBER: PCT/EP99/00696
PRIORITY FILING DATE: 1999-02-03
PRIORITY APPLICATION NUMBER: D 198 04 210.8
PRIORITY FILING DATE: 1998-02-03
NUMBER OF SEQ ID NOS: 41
SEQ ID NO 2
LENGTH: 255
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE: Artificial Sequence
OTHER INFORMATION: M1A-chain
FEATURE:
NAME/KEY: SITE
LOCATION: 15
OTHER INFORMATION: product= "Xaa is Asp or Glu"
OTHER INFORMATION: /label= Xaa1
FEATURE:
NAME/KEY: SITE
LOCATION: 63
OTHER INFORMATION: product= "Xaa is Gly or Gln"
OTHER INFORMATION: /label= Xaa2
FEATURE:
NAME/KEY: SITE
LOCATION: 66
OTHER INFORMATION: product= "Xaa is Ile or Val"
OTHER INFORMATION: /label= Xaa3
FEATURE:
NAME/KEY: SITE
LOCATION: 75
OTHER INFORMATION: product= "Xaa is Leu or Ala"
OTHER INFORMATION: /label= Xaa4
FEATURE:
NAME/KEY: SITE
LOCATION: 107
OTHER INFORMATION: product= "Xaa is missing"
OTHER INFORMATION: /label= Xaa5
FEATURE:
NAME/KEY: SITE
LOCATION: 113
OTHER INFORMATION: product= "Xaa is Asn or Thr"
OTHER INFORMATION: /label= Xaa6
FEATURE:
NAME/KEY: SITE
LOCATION: 117
OTHER INFORMATION: product= "Xaa is Pro or Thr"
OTHER INFORMATION: /label= Xaa7
FEATURE:
NAME/KEY: SITE
LOCATION: 134
OTHER INFORMATION: product= "Xaa is Asp or Glu"
OTHER INFORMATION: /label= Xaa8
FEATURE:
NAME/KEY: SITE
LOCATION: 141
OTHER INFORMATION: product= "Xaa is Ser or Thr"
OTHER INFORMATION: /label= Xaa9
FEATURE:
NAME/KEY: SITE
LOCATION: 145
OTHER INFORMATION: product= "Xaa is Phe or Tyr"
OTHER INFORMATION: /label= Xaa10
FEATURE:
NAME/KEY: SITE
LOCATION: 152
OTHER INFORMATION: product= "Xaa is Thr or Ala"
OTHER INFORMATION: /label= Xaa11
FEATURE:
NAME/KEY: SITE
LOCATION: 177

OTHER INFORMATION: product= "Xaa is Ala or Tyr"
OTHER INFORMATION: /label= Xaa12
FEATURE:
NAME/KEY: SITE
LOCATION: 180
OTHER INFORMATION: product= "Xaa is Tyr or Asp"
OTHER INFORMATION: /label= Xaa13
FEATURE:
NAME/KEY: SITE
LOCATION: 185
OTHER INFORMATION: product= "Xaa is Ala or Glu"
OTHER INFORMATION: /label= Xaa14
FEATURE:
NAME/KEY: SITE
LOCATION: 191
OTHER INFORMATION: product= "Xaa is Val or Met"
OTHER INFORMATION: /label= Xaa15
FEATURE:
NAME/KEY: SITE
LOCATION: 219
OTHER INFORMATION: product= "Xaa is Ile or Phe"
OTHER INFORMATION: /label= Xaa16
FEATURE:
NAME/KEY: SITE
LOCATION: 224
OTHER INFORMATION: product= "Xaa is Pro or Ser"
OTHER INFORMATION: /label= Xaa17
FEATURE:
NAME/KEY: SITE
LOCATION: 225
OTHER INFORMATION: product= "Xaa is Pro or Thr"
OTHER INFORMATION: /label= Xaa18
FEATURE:
NAME/KEY: SITE
LOCATION: 232
OTHER INFORMATION: product= "Xaa is Thr or Ser"
OTHER INFORMATION: /label= Xaa19
FEATURE:
NAME/KEY: SITE
LOCATION: 236
OTHER INFORMATION: product= "Xaa is Asp or Ser"
OTHER INFORMATION: /label= Xaa20
US-09-601-667B-2

alignment scores: Length: 88
Quality: 304.00
Ratio: 4.222
Percent Similarity: 81.818 Percent Identity: 69.318

alignment_block:
US-09-627-165B-13 x US-09-601-667B-2 ..
Align seg 1/1 to: US-09-601-667B-2 from: 1 to: 255

1 GCCAGATTCATCCCATGTCGAGGCTTCCGCCGAATTAACATGG 50
|||||
168 Alargphehnsnproileutrparg**ArgGln**IleasnSerg1 184
51 GGAATCTTCTCCACCAACATGTAATGCTCGAGCTGACAGAGTGGG 100
|||||
184 y**SerPheleuproasp**TyrmetleugluuulgutrhserTPG 201
101 GTGCAATCCACCAAGTCCAGCAGTCCMAAGATGACATTTTAATAC 150
|||||
201 TgInginserthrglnvalgInhiserThraspelyalPheasn 217
151 CAATTAAGATTGCAGATTTCGGCGGTAATTTGTGACGNTGACAACT 200
|||||
218 Pro**ArgleuAlaIle*****GlyAsnPhelvalThleu**AsnVa 234
201 TCCGACGTGATCTCAGCTTGGCGATCATGTTGTTCGAATGCAAGTGC 250
|||||

234 lArg***ValIleAlaSerLeuAlaIleMetLeuPheValCysGlyGluA 251
 251 GGGCATTCCTCCT 264
 ||||| |||||
 251 rProSerSerSer 255

seq_name: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:US-09-601-667B-41

seq_documentation_block:

Sequence 41, Application US/09601667B
 GENERAL INFORMATION:
 APPLICANT: Morris, Peter
 APPLICANT: Stiefel, Thomas
 APPLICANT: Voelker, Wolfgang
 APPLICANT: Wolters, Peter
 TITLE OR INVENTION: Recombinant Mistletoe Lectins
 FILE REFERENCE: 29841/36636
 CURRENT APPLICATION NUMBER: US/09/601,667B
 PRIOR FILING DATE: 2000-10-06
 PRIOR APPLICATION NUMBER: PCT/EP99/00696
 PRIOR FILING DATE: 1999-02-03
 PRIOR APPLICATION NUMBER: D 198 04 210.8
 NUMBER OF SEQ ID NOS: 41
 SEQ ID NO 41
 LENGTH: 256
 TYPE: PRP
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: MLA-chain
 NAME/KEY: SITE
 LOCATION: (15)..(15)
 OTHER INFORMATION: Xaa is Asp or Glu
 FEATURE:
 NAME/KEY: SITE
 LOCATION: (63)..(63)
 OTHER INFORMATION: Xaa is Gly or Glu
 FEATURE:
 NAME/KEY: SITE
 LOCATION: (66)..(66)
 OTHER INFORMATION: Xaa is Ile or Val
 FEATURE:
 NAME/KEY: SITE
 LOCATION: (75)..(75)
 OTHER INFORMATION: Xaa is Leu or Ala
 FEATURE:
 NAME/KEY: SITE
 LOCATION: (114)..(114)
 OTHER INFORMATION: Xaa is Asn or Thr
 FEATURE:
 NAME/KEY: SITE
 LOCATION: (118)..(118)
 OTHER INFORMATION: Xaa is Pro or Thr
 FEATURE:
 NAME/KEY: SITE
 LOCATION: (135)..(135)
 OTHER INFORMATION: Xaa is Asp or Glu
 FEATURE:
 NAME/KEY: SITE
 LOCATION: (142)..(142)
 OTHER INFORMATION: Xaa is Ser or Thr
 FEATURE:
 NAME/KEY: SITE
 LOCATION: (146)..(146)
 OTHER INFORMATION: Xaa is Phe or Tyr
 FEATURE:
 NAME/KEY: SITE
 LOCATION: (153)..(153)
 OTHER INFORMATION: Xaa is Thr or Ala
 FEATURE:
 NAME/KEY: SITE
 LOCATION: (178)..(178)

OTHER INFORMATION: Xaa is Ala or Tyr
 FEATURE:
 NAME/KEY: SITE
 LOCATION: (181)..(181)
 OTHER INFORMATION: Xaa is Tyr or Asp
 FEATURE:
 NAME/KEY: SITE
 LOCATION: (186)..(186)
 OTHER INFORMATION: Xaa is Ala or Glu
 FEATURE:
 NAME/KEY: SITE
 LOCATION: (192)..(192)
 OTHER INFORMATION: Xaa is Val or Met
 FEATURE:
 NAME/KEY: SITE
 LOCATION: (220)..(220)
 OTHER INFORMATION: Xaa is Ile or Phe
 FEATURE:
 NAME/KEY: SITE
 LOCATION: (225)..(225)
 OTHER INFORMATION: Xaa is Pro or Ser
 FEATURE:
 NAME/KEY: SITE
 LOCATION: (226)..(226)
 OTHER INFORMATION: Xaa is Pro or Thr
 FEATURE:
 NAME/KEY: SITE
 LOCATION: (233)..(233)
 OTHER INFORMATION: Xaa is Thr or Ser
 FEATURE:
 NAME/KEY: SITE
 LOCATION: (237)..(237)
 OTHER INFORMATION: Xaa is Asp or Ser
 US-09-601-667B-41

alignment_scores:
 Quality: 304.00 Length: 88
 Ratio: 4.222 Gaps: 0
 Percent Similarity: 81.818 Percent Identity: 69.318

alignment_block:

US-09-627-165b-13 x US-09-601-667B-41 ..

Align seg 1/1 to: US-09-601-667B-41 from: 1 to: 256

1 GCCAGATTCATCCATCCTGTCGAGGCTTCGCCGCAAAATTAACAGTGG 50
 |||||
 169 AlaArgPheAsnProIleLeuTrpArg***ArgGln***IleAsnSerG 185
 51 GGAGTCNTCTCCACCAACATGTCATGCTGAGCTGAGCAGAGCTTGGG 100
 | ||| |||||
 185 y***SerPheLeuProasp***TyrMetLeuGlnLeuGlnThrSerTrp 202
 101 GTCCGACATCCACCAACATGTCAGCAGTCAAGATGTCATTTTAATAC 150
 |||||
 202 lYgInGlnSerThrGlnValGlnHisSerThrAspGlyValPheAsn 218
 151 CAAATTAAGATTGCAGATTTCGCCGCGTAACCTTGTACGNTGACATAT 200
 ||||| |||||
 219 Pro***ArgLeuAlaIle*****GlyAsnPheValThrLeu***Asn 235
 201 TCCGACAGTGCATCCAGCTTGGCGATCATGTTGTAATGACAGTGTG 250
 ||||| |||||
 235 lArg***ValIleAlaSerLeuAlaIleMetLeuPheValCysGlyGlu 252
 251 GGGCATTCCTCCT 264
 ||||| |||||
 252 rProSerSerSer 256

seq_name: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep:US-10-083-336A-1
 seq_documentation_block:

CORRESPONDENCE ADDRESS:

TCGCGACGTGATCTCCAGCTTGGCGATCATGTTGT

245 IsertlleuileProilleleAlaLeuMetValIyrargCysAlaProp 262
 251 GGCCATTCGC 261
 262 roProSerSer 265

seq_name: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep:US-10-127-890-7

seq_documentation_block:

Sequence 7, Application US/10127890

GENERAL INFORMATION:

APPLICANT: Better, Marc D.

Studnika, Gary M.

TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating

Proteins

NUMBER OF SEQUENCES: 173

CORRESPONDENCE ADDRESSES:

ADDRESSEE: McAndrews, Held & Malloy, Ltd.

STREET: 500 West Madison Street, 34th floor

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60661

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/127,890

FILING DATE: 23-Apr-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/646,360

FILING DATE: 13-May-1996

APPLICATION NUMBER: PCT/US94/05348

FILING DATE: 12-May-1994

APPLICATION NUMBER: US 08/064,691

FILING DATE: 12-May-1993

APPLICATION NUMBER: US 07/988,430

FILING DATE: 09-DEC-1992

APPLICATION NUMBER: US 07/901,707

FILING DATE: 19-JUN-1992

APPLICATION NUMBER: US 07/787,567

FILING DATE: 04-NOV-1991

ATTORNEY/AGENT INFORMATION:

NAME: McNicholas, Janet M.

REGISTRATION NUMBER: 32,918

REFERENCE/DOCKET NUMBER: 200-70.P4

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/707-8889

TELEFAX: 312/707-9155

INFORMATION FOR SEQ ID NO: 7:

LENGTH: 263 amino acids

SEQUENCE CHARACTERISTICS:

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 7:

US-10-127-890-7

alignment_scores:

Quality: 92.00 Length: 82

Ratio: 1.769 Gaps: 3

Percent Similarity: 63.415 Percent Identity: 34.146

alignment_block:

US-09-627-165b-13 x US-10-127-890-7 ..

Align seg 1/1 to: US-10-127-890-7 from: 1 to: 263

1 GCCAGATTCATCCATCCTGAGGAGCTTGCGCGCAATTACAGTCG 50
 162 AAtargPhelystyrlleGlugInleInleInleValatYrargAs 178
 51 GGAGTCNTCTCCACCAACATGTACATGCTGAGCTGAGACGAGTTGGG 100
 178 pglu.....ValProSerLeuAlaThrIleSerLeuGluAsnSerTrps 193
 101 GTGCACATTCACCCCAAGTCAG.....CAGTCACAGATGCAATTTT 144
 193 ergIleuSerIysGlnIleGlnleuAlaGlnIysAsnngIyIlePhe 209
 145 AATACCCAAATATAGATTCGAGATTTCCGCGGTACTTGTGACGNTGAG 194
 210 ArghrProIleValleuValAspAsnIysGlyAsnArgValGlnIleTh 226
 195 CAATGTT.....CGCGAGCTGATCTCCAGCTTGGCGGATCATCTTG 234
 226 rAsnValThrSerIysValIleThrSerAsnIleGlnleuLeu 241

seq_name: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep:US-10-127-890-9

seq_documentation_block:

Sequence 9, Application US/10127890

GENERAL INFORMATION:

APPLICANT: Better, Marc D.

Studnika, Gary M.

TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating

Proteins

NUMBER OF SEQUENCES: 173

CORRESPONDENCE ADDRESSES:

ADDRESSEE: McAndrews, Held & Malloy, Ltd.

STREET: 500 West Madison Street, 34th floor

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60661

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/127,890

FILING DATE: 23-Apr-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/646,360

FILING DATE: 13-May-1996

APPLICATION NUMBER: PCT/US94/05348

FILING DATE: 12-May-1994

APPLICATION NUMBER: US 08/064,691

FILING DATE: 12-May-1993

APPLICATION NUMBER: US 07/988,430

FILING DATE: 09-DEC-1992

APPLICATION NUMBER: US 07/901,707

FILING DATE: 19-JUN-1992

APPLICATION NUMBER: US 07/787,567

FILING DATE: 04-NOV-1991

ATTORNEY/AGENT INFORMATION:

NAME: McNicholas, Janet M.

REGISTRATION NUMBER: 32,918

REFERENCE/DOCKET NUMBER: 200-70.P4

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/707-8889

TELEFAX: 312/707-9155

INFORMATION FOR SEQ ID NO: 9:

LENGTH: 261 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-10-127-890-9

alignment_scores:
Quality: 79.00 Length: 59
Ratio: 2.026 Gaps: 1
Percent Similarity: 66.102 Percent Identity: 33.698

alignment_block:

US-09-627-165b-13 x US-10-127-890-9 ..

Align seg 1/1 to: US-10-127-890-9 from: 1 to: 261

1 GCCAGATTCATCCCATGTCGAGCGCTCGCCGCAATTAACAGTGG 50
|||||
177 ALEATGPHETLTYRILEGLUANGINVALYSTRASNPHEASNTARGAS 193
51 GGAGTCNTCTCCACCAACATGTACATGCTGCGAGTCGAGACGAGTGGG 100
|||
193 pPheSer.....ProSnpAspLysValLeuAspLeuGluAsnTrpG 208
101 GTGACATCCACCCAGTCAGCATGTCACAGATGGCATTTTAAATACC 150
|||
208 LysLysLeuSerThrAlaLeuHisAsnSerLysAsnLeuValLeuProLys 224
151 CAAATAGATGTCAGATTTCCGCCGGT 177
|||||
225 ProLeuGluLeuLysAsnAlaAspGly 233

seq_name: /cgn2_6/prodata/1/paa/US10_NEW_COMB.pep:US-10-127-890-102

seq_documentation_block:

Sequence 102, Application US/10127890

GENERAL INFORMATION:

APPLICANT: Better, Marc D.

Carroll, Stephen F.

Studenka, Gary M.

TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating

Proteins

NUMBER OF SEQUENCES: 173

CORRESPONDENCE ADDRESS:

ADDRESSEE: McAndrews, Held & Malloy, Ltd.

STREET: 500 West Madison Street, 34th floor

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60661

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/127,890

FILING DATE: 23-Apr-2002

CLASSIFICATION: <Unknown>

Prior APPLICATION DATA:

APPLICATION NUMBER: US/08/646,360

FILING DATE: 13-May-1996

APPLICATION NUMBER: PCT/US94/05348

FILING DATE: 12-May-1994

APPLICATION NUMBER: US 08/064,691

FILING DATE: 12-May-1993

APPLICATION NUMBER: US 07/988,430

FILING DATE: 09-Dec-1992

APPLICATION NUMBER: US 07/901,707

FILING DATE: 19-JUN-1992

APPLICATION NUMBER: US 07/787,567

FILING DATE: 04-NOV-1991

ATTORNEY/AGENT INFORMATION:

NAME: McNicholas, Janet M.

REGISTRATION NUMBER: 32,918

REFERENCE/DOCKET NUMBER: 200-70-.p4

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/707-9155

TELEFAX: 312/707-8889

TELEX: 650 388-1248

INFORMATION FOR SEQ ID NO: 102:

SEQUENCE CHARACTERISTICS:

LENGTH: 251 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 102:

US-10-127-890-102

alignment_scores:

Quality: 75.50 Length: 82

Ratio: 1.541 Gaps: 2

Percent Similarity: 59.756 Percent Identity: 28.049

alignment_block:

US-09-627-165b-13 x US-10-127-890-102 ..

Align seg 1/1 to: US-10-127-890-102 from: 1 to: 251

1 GCCAGATTCATCCCATGTCGAGCGCTCGCCGCAATTAACAGTGG 50
|||||
168 ALEATGPHETLTYRILEGLUANGINLEIARGASN.....AsnProG1 182
51 GGAGTCNTCTCCACCAACATGTACATGCTGAGTCGAGACGAGTGGG 100
|||
182 nGlnArgLysLeuArgProAlaAsnAsnThrLysSerLeuGluAsnLysTrpG 199
101 GTGACATCCACCCAGTCAGCATGTCACAGATGGCATTTTAAAT 147
|||
199 LysLysLeuSerPheGlnIleArgThrSerGlyAlaAsnGlyMetPheSer 215
148 ACCCAATAGATTCAGATTTCCGCCGGTAACTTGTGACGNTGACCAA 197
|||||
216 GlnAlaValGluLeuGluArgAlaAsnGlyLysTyTyValThrAl 232
198 TGTTCGCGAGCTGATCTCCAGCTTGCGATCATGTTGTTCAATGC 243
|||
232 aValAspGlnValLysProLysIleAlaLeuLeuLysPheValCys 247

seq_name: /cgn2_6/prodata/1/paa/US10_NEW_COMB.pep:US-10-127-890-111

seq_documentation_block:

Sequence 111, Application US/10127890

GENERAL INFORMATION:

APPLICANT: Better, Marc D.

Carroll, Stephen F.

Studenka, Gary M.

TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating

Proteins

NUMBER OF SEQUENCES: 173

CORRESPONDENCE ADDRESS:

ADDRESSEE: McAndrews, Held & Malloy, Ltd.

STREET: 500 West Madison Street, 34th floor

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60661

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/127,890

FILING DATE: 23-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/646,360
FILING DATE: 13-MAY-1996
APPLICATION NUMBER: PCT/US94/05348
FILING DATE: 12-MAY-1994
APPLICATION NUMBER: US 08/064,691
FILING DATE: 12-MAY-1993
APPLICATION NUMBER: US 07/988,430
FILING DATE: 09-DEC-1992
APPLICATION NUMBER: US 07/901,707
FILING DATE: 19-JUN-1992
APPLICATION NUMBER: US 07/787,567
FILING DATE: 04-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: McNicholas, Janet M.
REGISTRATION NUMBER: 32,918
REFERENCE/DOCKET NUMBER: 200-70.P4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/707-8889
TELEFAX: 312/707-9155
TELEX: 650 368-1248
INFORMATION FOR SEQ ID NO: 111:
SEQUENCE CHARACTERISTICS:
LENGTH: 251 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 111:
US-10-127-890-111

alignment_scores:
Quality: 75.50 Length: 82
Ratio: 1.541 Gaps: 2
Percent Similarity: 59.756 Percent Identity: 28.049

alignment_block:

US-09-627-165b-13 x US-10-127-890-111 ..

Align seg 1/1 to: US-10-127-890-111 from: 1 to: 251

```
1 GCCGATTCAATCCGATCCTGCGAGCGCTGCGCGCAATTAAACAGTGG 50
|||||
168 AlaArgPheThrPheIleGluAsnGlnIleArgAsn.....AsnPheG1 182
51 GGAGTCNTCTCCACCAACATGTACATGCTCGAGCTGAGACGAGTGGG 100
|||||
182 nglnArgIleArgProAlaAsnAsnThrIleSerLeuGluAsnLysTrpG 199
101 GTGCGAATCCGACCAAGTCGACGATGC...AAGATGGCATTTTAAAT 147
|||||
199 LysIleuSerPheGlnIleArgThrSerGlyAlaAsnIleuMetPheSer 215
148 ACCCAATATAGATTCGAGATTTCCGCCGGTAACTTGTGACGNTGACGAA 197
|||||
216 GluAlaValGluLeuGluArgAlaAsnGlyLysIleValTrpAl 232
198 TGTTCGCGAGTGCATCTCCAGCTTGCGGATCATGTGTTCGATATGC 243
|||||
232 aValAspGlnValLysProLysIleAlaLeuIleuLysPheValCys 247
```


seq_documentation_block:

```
Sequence 8, Application US/09347064
GENERAL INFORMATION:
APPLICANT: Eck, Jurgen
APPLICANT: Schmidt, Arno
APPLICANT: Linke, Holger
TITLE OF INVENTION: Recombinant Fusion Proteins Based on
TITLE OF INVENTION: Ribosome-Inactivating Proteins of the mistletoe Viscum
FILE REFERENCE: 09282-5
CURRENT APPLICATION NUMBER: US/09/347,064
CURRENT FILING DATE: 1999-07-02
EARLIER APPLICATION NUMBER: PCT/EP98/00009
EARLIER FILING DATE: 1998-01-02
EARLIER APPLICATION NUMBER: EP 97 10 0012.0
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 8
LENGTH: 252
TYPE: PRT
ORGANISM: Viscum album
US-09-347-064-8
```

alignment_scores:

```
Quality: 619.00 Length: 138
Ratio: 4.654 Gaps: 0
Percent Similarity: 96.377 Percent Identity: 86.232
```

alignment_block:

```
US-09-627-165B-15 x US-09-347-064-8 ..
```

```
Align seg 1/1 to: US-09-347-064-8 from: 1 to: 252
```

```
1 TACACGATCTGGAGCGATACCGCGTCATAGGACCAATCCCTCTGG 50
|||||
115 TyrProaspLeuGlnArgTyrAlaGlyHisArgAspGlnIleProLeuG 131
51 TATAGAGAACTCAATCCGTCCTCGCGCTTCGTTATCCAGCGCGCA 100
|||||
131 YtleaspGlnIleuIleGlnSerValThrAlaLeuArgPheProGlyG 148
101 GCACCGGGGCCCAAGCTCGTCCCTTATATCCTCATTCAGATGATCC 150
|||||
148 erThrArgThrGlnAlaArgSerIleLeuIleLeuIleGlnMetIleSer 164
151 GAGCGCGGAGATTCATCCGATCTTTGGAGGCGTCGCAATACATTA 200
|||||
165 GlnAlaAlaArgPheAsnProIleLeuTrpArgAlaArgGlnTyrIleAs 181
201 CAGCGGGGAGTCATTTCTCCGACATGTCATGCTCCAGCTGGAGACTA 250
|||||
181 nSerGlyAlaSerPheLeuProAspValTyrMetLeuGlnLeuGluThrS 198
251 GTTGGGGCAACAATCCAGCAAGTCGACAGCTTACGGATCGCGTTT 300
|||||
198 erTrpGlyGlnGlnSerThrGlnValGlnHisSerThrAspGlyValPhe 214
301 AATAACCATTTGGGTGGTATATCCACCGGTAATCTGTACACTGAG 350
|||||
215 AsnAsnProIleArgLeuAlaIleProProGlyAsnPheValThrLeuTh 231
351 CAATGTCGCGAGTCGATCGCCAGCTTAGGATCATGCTGTTGATGTA 400
|||||
231 rAsnValAlaArgAspValIleAlaSerLeuAlaIleMetLeuPheValCysG 248
401 GGGACCGACCATCT 414
|||||
248 YglnArgProSer 252
```

```
seq_name: /cgn2_6/ptodata/2/paa/US093_COMB.pep:US-09-347-064-8
```

seq_documentation_block:

```
Sequence 8, Application US/09347064A
GENERAL INFORMATION:
APPLICANT: Eck, Jurgen
APPLICANT: Schmidt, Arno
APPLICANT: Linke, Holger
TITLE OF INVENTION: Recombinant Fusion Proteins Based on
TITLE OF INVENTION: Ribosome-Inactivating Proteins of the mistletoe Viscum
FILE REFERENCE: 09282-5
CURRENT APPLICATION NUMBER: US/09/347,064A
CURRENT FILING DATE: 1999-07-02
EARLIER APPLICATION NUMBER: PCT/EP98/00009
EARLIER FILING DATE: 1998-01-02
EARLIER APPLICATION NUMBER: EP 97 10 0012.0
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 8
LENGTH: 252
TYPE: PRT
ORGANISM: Viscum album
US-09-347-064-8
```

alignment_scores:

```
Quality: 619.00 Length: 138
Ratio: 4.654 Gaps: 0
Percent Similarity: 96.377 Percent Identity: 86.232
```

alignment_block:

```
US-09-627-165B-15 x US-09-347-064-8 ..
```

```
Align seg 1/1 to: US-09-347-064-8 from: 1 to: 252
```

```
1 TACACGATCTGGAGCGATACCGCGTCATAGGACCAATCCCTCTGG 50
|||||
115 TyrProaspLeuGlnArgTyrAlaGlyHisArgAspGlnIleProLeuG 131
51 TATAGAGAACTCAATCCGTCCTCGCGCTTCGTTATCCAGCGCGCA 100
|||||
131 YtleaspGlnIleuIleGlnSerValThrAlaLeuArgPheProGlyG 148
101 GCACCGGGGCCCAAGCTCGTCCCTTATATCCTCATTCAGATGATCC 150
|||||
148 erThrArgThrGlnAlaArgSerIleLeuIleLeuIleGlnMetIleSer 164
151 GAGCGCGGAGATTCATCCGATCTTTGGAGGCGTCGCAATACATTA 200
|||||
165 GlnAlaAlaArgPheAsnProIleLeuTrpArgAlaArgGlnTyrIleAs 181
201 CAGCGGGGAGTCATTTCTCCGACATGTCATGCTCCAGCTGGAGACTA 250
|||||
181 nSerGlyAlaSerPheLeuProAspValTyrMetLeuGlnLeuGluThrS 198
251 GTTGGGGCAACAATCCAGCAAGTCGACAGCTTACGGATCGCGTTT 300
|||||
198 erTrpGlyGlnGlnSerThrGlnValGlnHisSerThrAspGlyValPhe 214
301 AATAACCATTTGGGTGGTATATCCACCGGTAATCTGTACACTGAG 350
|||||
215 AsnAsnProIleArgLeuAlaIleProProGlyAsnPheValThrLeuTh 231
351 CAATGTCGCGAGTCGATCGCCAGCTTAGGATCATGCTGTTGATGTA 400
|||||
231 rAsnValAlaArgAspValIleAlaSerLeuAlaIleMetLeuPheValCysG 248
401 GGGACCGACCATCT 414
|||||
248 YglnArgProSer 252
```

```
seq_name: /cgn2_6/ptodata/2/paa/US093_COMB.pep:US-09-347-064-2
```

```

seq.documentation_block:
: Sequence 2, Application US/09347064A
: GENERAL INFORMATION:
: APPLICANT: Eck, Jurgen
: APPLICANT: Schmidt, Arno
: APPLICANT: Zinke, Holger
: TITLE OF INVENTION: Recombinant Fusion Proteins Based on
: TITLE OF INVENTION: Ribosome-Inactivating Proteins of the mistletoe viscum
: TITLE OF INVENTION: album
: FILE REFERENCE: 09282-5
: CURRENT APPLICATION NUMBER: US/09/347,064A
: CURRENT FILING DATE: 1999-07-02
: EARLIER APPLICATION NUMBER: PCT/EP98/00009
: EARLIER FILING DATE: 1998-01-02
: EARLIER APPLICATION NUMBER: EP 97 10 0012.0
: EARLIER FILING DATE: 1997-01-02
: NUMBER OF SEQ ID NOS: 38
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 2
: LENGTH: 252
: TYPE: prt
: ORGANISM: Viscum album
:
US-09-347-064-2

```

alignment_scores:	length:	137
Quality:	Gaps:	0
Ratio:	Percent Identity:	86.131
Percent Similarity:		96.350

alignment_block: US-09-347-064-2

Align seg 1/1 to: US-09-347-064-2 from: 1 to: 252

1 TACACAGATCTGGAGCGATACACGCCGCTATACGAGACAGATCCCTTGGG 50
111
116 TyrProAspLeuGlnGlyTyrAlaGlyHisArgSerGlnIleProLeuG1 132
51 TATAGAGCAATCAATCAATCCGTCGCGCGCTCGTTATCCAGCGGCA 100
113 yIleAspIleuIleGlnSerValThrIleLeuArgPheProGlyGlyS 149
101 GCACCCGGGCCCAAGCTCGTTCCTTATATCTCATATGATGATCC 150
149 eArThArgThGlnAlaIArgSerIleuIleuIleuIleGlnIleSer 165
151 GAGGCGCGGAGATCAATCAATCCCATCTTTGGAGGGCTCGCCATACATTAA 200
166 GlnAlaIleAlaArgPheAsnProIleuLeuTrpArgAlaArgIleuTrpLeas 182
201 CAGCGGGGAGATCTATTTCTTCCCGACAGTACATCCCTGAGCTGGAAGCTA 250
182 nSerGlyAlaSerPheLeuProAspValTyrMetLeuGlnLeuGlnThrs 199
251 GTTGGGGCCCAAAATGCCAGCAGATGCCAGCTACCGATGCCGTTTTT 300
199 eTrpIleGlnIleSerThrGlnValGlnHisSerThrAspGlyValPhe 215
301 AATTAACCATTTGGTGGGTATATTCACCGGTAACTCTGAGAGCTTGAG 350
216 AsnAsnIleTrpLeuArgLeuAlaIleProProGlyAsnPheValThrLeuTh 232
351 CAATGTTCCGAGCGATGATCCGACGCTTAACGATCATGTTGTTTATATGTA 400
232 rAsnValArgAspValIleIleIleSerLeuAlaIleMetLeuPheValCysG 249
401 GGGAGCGACCA 411
249 IyGlnArgPro 252

seq_name: /cgn2_6/ptodata/2/paa/US096_COMB.pep:US-09-627-165B-4

seq_documentation_block:

```
; Sequence 4, Application US/09627165B
; GENERAL INFORMATION:
; APPLICANT: KIM, Jong-Bae
; TITLE OF INVENTION: CRUDE EXTRACT FROM VISCUM album coloratum, AND PROTEINS
; FILE REFERENCE: AND LECTINS ISOLATED THEREFROM
; CURRENT APPLICATION NUMBER: US/09/627,165B
; CURRENT FILING DATE: 2000-07-27
; NUMBER OF SEQ ID NOS: 16
; SEQ ID NO 4
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Viscum album coloratum
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 240
; OTHER INFORMATION: Xaa = any amino acid
US-09-627-165B-4
```

alignment_scores:

Quality:	614.00	Length:	140
Ratio:	4.582	Gaps:	0
Percent Similarity:	95.714	Percent Identity:	86.429

alignment_block:

US-09-627-165B-15 x US-09-627-165B-4 ..

Align seg 1/1 to: US-09-627-165B-4 from: 1 to: 254

```
1 TACACAGATCTGAGAGCATACGCCGGTCATAGAGACAGATCCCTCTGG 50
115 TYRPROASPLEUGLARGTYRALAGLYHISARGASPRGIIIEPRLG 131
51 TATPAGAGACTCATTCATCCGTCGTCGAGGCTGTATCCAGAGCGGCA 100
131 YLIEASPRINLEUIEGLINSEVALSERIALLEUATGPHEPLOGLYSE 148
101 GCACCCGGGCCAGACGTCGTCCTTAAATCCATTCAGATGATCTCC 150
148 srtthArYaglnlaArgserPhelelleuIlleGlmetlleSer 164
151 GAGCCCGGAGATTCATCCATCTTTTGGAGGCTCGCAATACATTAA 200
165 GIUALAAlaArgpheaSnProIleuTrpArgAlaArgIntYrIleSe 181
201 CAGCGGGAGATCATTCCTCCGACATGACATGCTGAGAGCTGA 250
181 rserdlYglserPheleuProAspHrTYrIleuGlneugIntYrIleSe 198
251 GTTGGGGCCAAATCCACGCAAGTCACAGCTTACGAGATGGCTTTT 300
198 eTrtPdlYglInlserThrlGlnValGlnHisserThraspLYalPhe 214
301 AATAACCATTTGCGTGGTATATCCACCGGTAACCTTCGACGCTTGA 350
215 AsnaSnProIleArgleuThrlIleSerThrlGlyValPheValThrlSe 231
351 CAATGTCGACGATGCGCAGCTTAGCATGATCATGCTGTGTATGTA 400
231 rAsnValArgspValIleAlaSer**AlaIleMetleuPheValCysG 248
401 GGGACCGACCATCTCTCC 420
248 luhAspArgPserSerSer 254
```

seq_name: /cgn2_6/ptodata/2/paa/US096_COMB.pep:US-09-627-165B-6

seq_documentation_block:

```
; Sequence 6, Application US/09627165B
; GENERAL INFORMATION:
```

```
; APPLICANT: KIM, Jong-Bae
```

```
; TITLE OF INVENTION: CRUDE EXTRACT FROM VISCUM album coloratum, AND PROTEINS
```

```
; FILE REFERENCE: AND LECTINS ISOLATED THEREFROM
```

```
; CURRENT APPLICATION NUMBER: US/09/627,165B
```

```
; CURRENT FILING DATE: 2000-07-27
```

```
; NUMBER OF SEQ ID NOS: 16
```

```
; SEQ ID NO 6
```

```
; LENGTH: 256
```

```
; TYPE: PRT
```

```
; ORGANISM: Viscum album coloratum
```

```
; FEATURE:
```

```
; NAME/KEY: misc_feature
US-09-627-165B-6
```

alignment_scores:

Quality:	611.00	Length:	139
Ratio: <td>4.629<td>Gaps:<td>0</td></td></td>	4.629 <td>Gaps:<td>0</td></td>	Gaps: <td>0</td>	0
Percent Similarity: <td>94.964<td>Percent Identity:<td>86.331</td></td></td>	94.964 <td>Percent Identity:<td>86.331</td></td>	Percent Identity: <td>86.331</td>	86.331

alignment_block:

US-09-627-165B-15 x US-09-627-165B-6 ..

Align seg 1/1 to: US-09-627-165B-6 from: 1 to: 256

```
1 TACACAGATCTGAGAGCATACGCCGGTCATPAGAGACAGATCCCTCTGG 50
117 TYRALAAsPLeuGlLARGTYRALAGLYHISARGASPRGIIIEPRLG 133
51 TATPAGAGACTCATTCATCCGTCGTCGAGGCTGTATCCAGAGCGGCA 100
133 YARGlupProleuIleArgserValserAlaLeuAspYrProdlYglYs 150
101 GAGCCCGGCGCCAGCTCGTCGTCCTTAAATCCATTCAGATGATCTCC 150
150 eTrtthArYaglnlaArgserIlelleIleValIleGlmetIleSe 166
151 GAGCCCGGAGATTCATCCATCTTTTGGAGGCTCGCAATACATTAA 200
167 GIUALAAlaArgpheaSnProIleuTrpArgAlaArgIntYrIleAs 183
201 CAGCGGGAGATCATTCCTCCGACATGACATGCTGAGAGCTGA 250
183 nserdlYglserTYrleuProAspValYrMetleuGlneugIntYrIleSe 200
251 GTTGGGGCCAAATCCACGCAAGTCACAGCTTACGAGATGGCTTTT 300
200 eTrtPdlYglInlserThrlGlnValGlnHisserThraspLYalPhe 216
301 AATAACCATTTGCGTGGTATATCCACCGGTAACCTTCGACGCTTGA 350
217 AsnaSnProIleArgleuThrlIleSerThrlGlyAsnPheValTYrIleSe 233
351 CAATGTCGACGATGCGCAGCTTAGCATGATCATGCTGTGTATGTA 400
233 rAsnValArgspValIleAlaSerIleuGlYrIleMetValPheValCysA 250
401 GGGACCGACCATCTCTCC 417
250 rGAspArgSerSerSer 255
```

seq_name: /cgn2_6/ptodata/2/paa/US096_COMB.pep:US-09-627-165B-14

seq_documentation_block:

```
; Sequence 14, Application US/09627165B
; GENERAL INFORMATION:
; APPLICANT: KIM, Jong-Bae
; TITLE OF INVENTION: CRUDE EXTRACT FROM VISCUM album coloratum, AND PROTEINS
; FILE REFERENCE: AND LECTINS ISOLATED THEREFROM
; CURRENT APPLICATION NUMBER: US/09/627,165B
; CURRENT FILING DATE: 2000-07-27
```


Wed Jul 31 08:46:43 2002

us-09-627-165b-15.ram

```
;
; NUMBER OF SEQ ID NOS: 16
; SEQ ID NO 14
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Viscum album coloratum
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 7, 64, 111
; OTHER INFORMATION: Xaa - any amino acid
US-09-627-165B-14

alignment_scores:
  Quality: 408.50      Length: 122
  Ratio: 4.045         Gaps: 1
  Percent Similarity: 82.787   Percent Identity: 69.672

alignment_block:
US-09-627-165B-15 x US-09-627-165B-14 ..
Align seg 1/1 to: US-09-627-165B-14 from: 1 to: 119

157 GCGAGATTCATCCATCTTTGGAGGCGTCGCCAATACATTAACGCG 206
|||||
1 AlatrphsAnProIle**TtrpargleuArgrglnIleasnSergl 17
207 GGAGTCATTTCTCCGACATGTACATGCTCGAGCTGGAGACTAGTGGG 256
|||||
17 ygluSerSerProbsanMetlyrMeluGlulduGlnthrsertpg 34
257 GCCAACAATCCAGCAGTCCAGCAGTCTACGAGTCCGCTTTTATATAC 306
|||||
34 lYarGlnSerThnGlnValGlnGlnSerlyaspGlyIlephasnThr 50
307 CCATTCGCTGGGTATATCCACCGGTAACCTGTGACGCTGAGCAATGT 356
|||||
51 GlnIleArgLeuGlnIleSerAlaGlyaspheValThr**Sernsna 67
357 TCCCGACGATGCGCAGCTTACGATCATGTTGTTGATGTAGGAGAC 406
|||||
67 lArGaspValIleSerSerleuAlaIleMetleuPhelGlySerGlyA 84
407 GACCATCTCTCCGACGCTGCGCTATTCGCGCGTGCATACGACCGTC 456
|||||
84 rgrProphSerSerleuAspHisProSerProleuLeuLeuArgSerVal 100
457 TTGGAAATATAGCGCGCGCTGACGATGTATACCTGCTGCTGCAACC 506
|||||
101 Valasp.....AlaAlaAsnAspValThrCysThr**SergLupr 114
507 CACCGTGGCGATCGTA 522
|||||
114 othrValArGlyLeval 119

seq_name: /cgn2_6/ptodata/2/paa/US095_COMB.pep:US-09-538-873-3

seq_documentation_block:
; Sequence 3, Application US/0953873
; GENERAL INFORMATION:
; APPLICANT: VITETTA, ELLEN S.
; APPLICANT: GHETIE, VICTOR F.
; APPLICANT: SMALLSHAW, JOAN
; APPLICANT: BALUNA, ROXANA G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR AMELIORATING VASCULAR LEAK
; TITLE OF INVENTION: SYNDROME (VLS)
; FILE REFERENCE: US/09/538,873
; CURRENT FILING DATE: 2000-03-30
; EARLIER FILING DATE: 1999-03-30
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
```

```
;
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Abrus precatorius
; US-09-538-873-3

alignment_scores:
  Quality: 243.00      Length: 145
  Ratio: 2.505         Gaps: 3
  Percent Similarity: 66.897   Percent Identity: 40.000

alignment_block:
US-09-627-165B-15 x US-09-538-873-3 ..
Align seg 1/1 to: US-09-538-873-3 from: 1 to: 251

1 TACACAGATCTGGAGCATACGCGGTCAT...AGGACAGATCCCTCT 47
|||
113 tyrGlyaspLeuGlnArGlyrPalahNisGlnSerArGlnGlnIleProle 129
48 GGATATAGAGAACTCATTCATCCGCTCTGCGCGCTTCGTTATCCAGCG 97
|||||
129 uGlyLeuGlnAlaLeuThrNisGlyIleSerPhePheArGlySerGlyGlyA 146
98 GCACACCCGCGCCAGCTGCTCCCTTATATATCTTCATTCAGATGATC 147
|||||
146 snAspAsnGlnGluysAlaArGlyrleuIleValIleIleGlnMetVal 162
148 TCCGAGCGCGCGATTCATCCATCTTTGAGGCGCTGCCAATACAT 197
|||
163 AlaAlaAlaAlaArGlyrheArGlyrIleSerAsnArGlyArGlyAsnIle 179
198 TAACAGCGGGAGTCATTTCTCCGACATGTACATGCTGACGTGAGGA 247
|||||
179 eGlnThrIleYThrAlaPheGlnProAspAlaIleMetleuSerleuGlnA 196
248 CTAATGGGGGCCAACAAATCCACCAAGCTCCAGCAGCTACGAGTGCCT 297
|||||
196 snAsnTrpAspAsnleuSerArGlyValGlnIleuSerValGlnAspThr 212
298 TTTATATACCATTTGCTGGGTATATCCACCGGTAACCTGTGACACT 347
|||
213 PheProAsnGln.....ValThrIle 219
348 GACCAATGTTTCGCAC.....G 364
|||||
219 uThrAsnIleArGAsnGluProValIleValAspSerleuSerHisProT 236
365 TGATCCCGCAGCTTACGATCATGTTGTTGATGT 399
|||||
236 hrValAlaValIleuAlaLeuMetleuPheValCys 247

seq_name: /cgn2_6/ptodata/2/paa/US096_COMB.pep:US-09-668-419-3

seq_documentation_block:
; Sequence 3, Application US/09668419
; GENERAL INFORMATION:
; APPLICANT: VITETTA, ELLEN S.
; APPLICANT: GHETIE, VICTOR F.
; APPLICANT: SMALLSHAW, JOAN E.
; APPLICANT: BALUNA, ROXANA G.
; TITLE OF INVENTION: RIGIN A CHAIN MUTANTS LACKING ENZYMATIC ACTIVITY AS VACCINES
; TITLE OF INVENTION: AGAINST AEROSOLIZED RIGIN
; FILE REFERENCE: US/09/668,419
; CURRENT FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Abrus precatorius
; US-09-668-419-3
```

alignment_scores:
 Quality: 243.00 Length: 145
 Ratio: 2.505 Gaps: 3
 Percent Similarity: 66.897 Percent Identity: 40.000

alignment_block:
 US-09-627-165b-15 x US-09-668-419-3 ..

Align seg 1/1 to: US-09-668-419-3 from: 1 to: 251

```

1 TACACAGATCTGAGCAGATACGCCGGTCAT...AGGACAGATCCCTCT 47
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
113 TYGGLASPLeuGluArgTrpAlaHisGlnSerArgGlnGlnIleProle 129
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
48 GGGTATAGAGAACTCATTCATCCCTCTCGGCGCTTGTATCCAGCGC 97
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
129 uGlyLeuGlnAlaLeuThrHisGlyIleSerPhePheArgSerGlyVal 146
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
98 GCAGACCCGGCCAGCTCTGTCCTTATATCCATTCATGATGATC 147
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
146 snAspAsnGluGluValAlaArgThrLeuIleValIleIleGlnMetVal 162
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
148 TCCGAGCGCCGGAGATTCATCCCATCTTTGGAGGCGTCCCAATACAT 197
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
163 AAlaAlaAlaAlaArgPheArgTrpIleSerAsnArgValAlaValSerI 179
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
198 TACAGCGCGGAGTCATTCCTTCCGACATGATCATCTCGAGCTGAGA 247
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
179 eGlnThrGlyThrAlaPheGlnProAspAlaAlaMetIleSerLeuGlnA 196
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
248 CTAGTTGGGGCCCAACATCCACAGTCCAGTCCAGATCGAGTGGCGTT 297
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
196 snAsnTrpAspAsnLeuSerArgGlyValGlnGlnSerValGlnAspThr 212
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
298 TTTATATACCCATTCGTTGGGTATATCCACCGGTACTCTGTCGACGT 347
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
213 PheProAsnGln.....ValThrIle 219
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
348 GAGCAATGTTCCGAC.....G 364
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
219 uThrAsnIleArgAsnGluProValIleValAspSerLeuSerHisProT 236
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
365 TGAATGCCAGCTTAGCGATCATGTTGTTGATGTT 399
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
236 hValAlaValAlaLeuAlaLeuMetLeuPheValCys 247

```

seq_name: /cgn2.6/plodata/2/paa/US07_COMB.pep:US-07-787-567A-1

seq_documentation_block:

```

; Sequence 1, Application US/07787567A
; GENERAL INFORMATION:
; APPLICANT: Better, Marc D.
; APPLICANT: Carroll, Steve F.
; APPLICANT: Lane, Julie A.
; TITLE OF INVENTION: Materials Comprising and Methods of
; TITLE OF INVENTION: Preparation and Use for Ribosome-Inactivating Proteins
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza, 20 South Clark
; STREET: Street
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/787,567A
; FILING DATE: 19911104
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Gruber, Lewis S.
; REGISTRATION NUMBER: 30,060
; REFERENCE/DOCKET NUMBER: 27129/30545
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 267 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-787-567A-1

```

alignment_scores:
 Quality: 217.00 Length: 143
 Ratio: 2.309 Gaps: 2
 Percent Similarity: 65.734 Percent Identity: 39.860

alignment_block:
 US-09-627-165b-15 x US-07-787-567A-1 ..

Align seg 1/1 to: US-07-787-567A-1 from: 1 to: 267

```

1 TACACAGATCTGAGCAGATACGCCGGTCAT...AGGACAGATCCCTCT 47
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
123 TYRAspArgLeuGluGlnLeuAlaArgGlnAsnLeuArgGlnsnIleIle 139
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
48 GGGTATAGAGAACTCATTCATCCGTCCTCGGCGCTTGTAT..... 90
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
139 uGlyAsnGlyProLeuGlnIleValAlaLeuSerAlaLeuTrpTYrTYrSerT 156
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
91 ..CAAGCGCGACGACCCGGCCCAAGCTCTGTCCTTATATCCATTC 138
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
156 hGlyGlyThrGlnLeuProThrLeuAlaArgSerPheIleIleCysIle 172
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
139 CAGATGATCTCCGAGCGCGAGATTCATCCCATCTTTGGAGGCTCG 188
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
173 GlnMetIleSerGlnAlaAlaArgPheGlnTrpIleGluGlyGlnMetArg 189
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
189 CCAATACATTCACAGCGGAGTCATTCCTCCGACATGATCATGCTCG 238
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
189 gThrArgIleArgTrpAsnArgArgSerAlaProAspProSerValIleT 206
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
239 AGCTGAGACATGATGTTGGGCGCAACATCCAGCAGTCCAGTCCAGT 288
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
206 hLeuGlnAsnSerTrpGlyArgLeuSerThrAlaIleGlnGlnSerAsn 222
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
288 GATGCGCTTTTAATACCATTCATTCGTTGGGTATCCACCGGTAATT 338
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
223 GlnGlyAlaPheAlaSerProIleGlnLeuGlnArgArgGlnsnIleSerI 239
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
339 CGTGAAGTTAGCAATGTTGCGAGCTGATCCGACGATTCAGGATCATGT 388
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
239 sPheSerValTrpAspValSerIleLeuIleProIleIleAlaLeuMetV 256
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
389 TGTATTGATGAGGACGACCATCTTCC 417
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
256 aLYrArgCysAlaProProProSerSer 265

```

seq_name: /cgn2.6/plodata/2/paa/US084_COMB.pep:US-08-406-801-34

seq_documentation_block:

```

; Sequence 34, Application US/08406801
; GENERAL INFORMATION:
; APPLICANT: Wright, Andrew F.

```

APPLICANT: Blakey, David C.
APPLICANT: Filton, John E.
APPLICANT: Lindholm, Lelf
APPLICANT: Lindgren, Peter
APPLICANT: Holmgren, Jan
TITLE OF INVENTION: Conjugates
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cushman, Darby & Cushman
STREET: 1100 New York Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/406,801
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/908,269
FILING DATE: 02-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Kokulis, Paul N.
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: PNK/3893/96491/MLW
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3000
TELEFAX: 202-822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 267 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-406-801-34

alignment_scores:
Quality: 217.00 Length: 143
Ratio: 2.309 Gaps: 2
Percent Similarity: 65.734 Percent Identity: 39.860

alignment_block:

US-09-627-165B-15 x US-08-406-801-34 ..

Align seg 1/1 to: US-08-406-801-34 from: 1 to: 267

```
1 TACACAGATCTGAGCGGATACGCGGTCAAT...AGGACGACATCCCTCT 47
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
123 TyrAspArgLeuGlnLeuAlaGlyAsnLeuArgGluAsnIleGluLe 139
48 GGGTATAGAGAACTCATTCATCCGTCGCGCGCTGTTAT..... 90
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
139 uGlyAsnGlyProLeuGlnLeuAlaIleSerAlaLeuIleTyrTyrSer 156
91 ..CCAGCGCGACAGCCCGGCCCAAGCTGTTCCCTTATATCCCTCAT 138
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
156 hrgIyGlyThrGlnLeuProThrLeuAlaArgSerPheIleIleCysIle 172
139 CAGATGATCTCGAGCGCGGAGATTCATCCATCCCTTTGGAGGGCTCG 188
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
173 GlnMetIleSerGlnAlaAlaArgPheGlnIleGluGluGluMetArg 189
189 CCAATACATTAAACAGCGGGAGTCATTTCTCCCGACATGATCATGCTCG 238
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
189 gThrArgIleArgTyrAsnArgArgSerAlaProAspProSerValIle 206
```

```
239 AGCTGAGACTGAGTGGGCCCAACAATCCAGCGCACTCAGAGTCTACG 288
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
206 hreGluAsnSerTrpGlyArgLeuSerThrAlaIleGlnIleGlnIle 222
289 GATGCGGCTTTTAAATACCATTTTCGTTGGGTATATCCACCGGTA 338
::: ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
223 GlnGlyAlaPheAlaSerProIleGlnLeuGlnArgArgAsnGlySer 239
339 CGTACGTTGAGCAATGTTCCGAGCGATGATCCGACCTAGGACATCAT 388
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
239 spheSerValTyrAspValSerIleLeuIleProIleIleAlaLeuMet 256
389 TGTGTGATGTAGGACGACCATCTTCC 417
::: ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
256 alTyrArgCysAlaProProSerSer 265
```

seq_name: /cgn2_6/ptodata/2/paa/US095_COMB.pep:US-09-538-873-1

seq_documentation_block:
Sequence 1, Application US/09538873
GENERAL INFORMATION:
APPLICANT: VITETTA, ELLEN S.
APPLICANT: GHERIE, VICTOR F.
APPLICANT: SMALISHAW, JOAN
APPLICANT: BALUNA, ROXANA G.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR AMELIORATING VASCULAR LEAK
FILE REFERENCE: US05:603
CURRENT APPLICATION NUMBER: US/09/538,873
EARLIER FILING DATE: 2000-03-30
EARLIER APPLICATION NUMBER: 60/126,826
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 267
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-538-873-1

alignment_scores:
Quality: 217.00 Length: 143
Ratio: 2.309 Gaps: 2
Percent Similarity: 65.734 Percent Identity: 39.860

alignment_block:

US-09-627-165B-15 x US-09-538-873-1 ..

Align seg 1/1 to: US-09-538-873-1 from: 1 to: 267

```
1 TACACAGATCTGAGCGGATACGCGGTCAAT...AGGACGACATCCCTCT 47
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
123 TyrAspArgLeuGlnLeuAlaGlyAsnLeuArgGluAsnIleGluLe 139
48 GGGTATAGAGAACTCATTCATCCGTCGCGCGCTGTTAT..... 90
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
139 uGlyAsnGlyProLeuGlnLeuAlaIleSerAlaLeuIleTyrTyrSer 156
91 ..CCAGCGCGACAGCCCGGCCCAAGCTGTTCCCTTATATCCCTCAT 138
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
156 hrgIyGlyThrGlnLeuProThrLeuAlaArgSerPheIleIleCysIle 172
139 CAGATGATCTCGAGCGCGGAGATTCATCCATCCCTTTGGAGGGCTCG 188
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
173 GlnMetIleSerGlnAlaAlaArgPheGlnIleGluGluGluMetArg 189
189 CCAATACATTAAACAGCGGGAGTCATTTCTCCCGACATGATCATGCTCG 238
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
189 gThrArgIleArgTyrAsnArgArgSerAlaProAspProSerValIle 206
```

```

: LENGTH: 267 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-09-610-838-1

alignment_scores:
    Quality: 217.00      Length: 143
    Ratio: 2.309         Gaps: 2
    Percent Similarity: 65.734      Percent Identity: 39.860

Alignment_block:
US-09-627-165B-15 x US-09-610-838-1 ..

Align seg 1/1 to: US-09-610-838-1 from: 1 to: 267

1 TACACAGATCTGGAGCGATACGCCGTCAT..AGGACCAATCCCTCT 47
  ||| |||||::: |||||::: |||::: ||| ||
123 TyrspApgLeuGIuGInLeuAlaGIyAsnLeuArgIuAsnIleGlu 139
  ||| ||| :::: |||||::: |||
48 GGGTAAAGAGAACTCAATCCGTCCTGGCGCTTCGTTAT..... 90
  ||| ||| :::: |||||::: |||
139 uGIyAsnGIyProLeuGIuGInuAlaIleSerAlaLeuIYrIYrIYSerT 156
  ||| ||| :::: |||||::: |||
91 ..CCAGCGCGCAGCACCCGGGCCCAAGCTGTTCCCTTAATCTCAT 138
  ||| ||| :::: |||||::: |||
156 hrGIyGIYThrGIuLeuProThrLeuAlaArgSerPheIleIleCysIle 172
  ||| ||| :::: |||||::: |||
139 CAGATGATCTCCGAGCGCGAGATTCATCCATCCATCTTTGGAGGCGTCG 188
  ||||| ||||| ||||| |||||::: ||| :::: ||
173 GInuGIleSerGIuAlaAlaArgPheGIuIYrIleGIuGIyGIuMetArg 189
  ||||| ||||| ||||| |||||::: ||| :::: ||
189 CCAATATCAATPACACGCGGAGATCATTTCTTCCCGACATGTACATGCTTCG 238
  ||| ||| :::: |||||::: |||||
189 gThArgIleArgTYrAsnArgArgSerAlaProAspProSerValIleT 206
  ||| ||| :::: |||||::: |||||
239 AGCTGGAGACTAGTTGGGGCCCAATCCACGCAAGTCCAGCAGTGTACG 288
  |||||::: ||||| ||||| :::: |||||::: |||
206 hrLeuGIuAsnSerIYrPcIyArgIeuserThrAlaIleGIuIeuserAsn 222
  |||||::: ||||| ||||| :::: |||||::: |||
289 GATGGCGTTTAAATACCACTTCGTTGGGTATATCCACCGGTAAC 338
  ::|||::: ||| :::: |||||::: |||
223 GInGIyAlaPheAlaSerProIleGIuLeuGIuArgArgAsnGIySerIY 239
  |||||::: ||||| ||||| :::: |||||::: |||
239 CGTGAGCTTGAGCAAGTTCGCGACGTATGCCACGCTTAGCGATCANGT 388
  ::|||::: ||| :::: |||||::: |||
239 sPheSerValTYrAspAlaSerIleLeuIleProIleIleAlaLeuMetY 256
  |||||::: ||| ||||| |||||
389 TGTTTGTATGTAGGAGCGACCACTCTCC 417
  ::|||::: ||| ||||| |||||
256 aIYrArgCysAlaProProSerSer 265
  |||||::: ||| ||||| |||||

seq_name: /cgn2_6/plodacta/2/paa/US096.COMB.pep:US-09-668-419-1

seq_documentation_block:
: Sequence 1, Application US/09668419
: GENERAL INFORMATION:
: APPLICANT: VITETTA, ELLEN S.
: APPLICANT: GHETTE, VICTOR F.
: APPLICANT: SMALLSHAW, JOAN E.
: APPLICANT: BALDUNA, ROXANA G.
: TITLE OF INVENTION: RICIN A CHAIN MUTANTS LACKING ENZYMIC ACTIVITY AS VACCINES
: TITLE OF INVENTION: AGAINST AEROSOLIZED RICIN
: FILE REFERENCE: UTSD:72305
: CURRENT APPLICATION NUMBER: US/09/668,419
: NUMBER OF SEQ ID NOS: 19
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 1
: LENGTH: 267
: TYPE: PRT
: ORGANISM: Artificial Sequence

```

[illegible]

Align seg 1/1 to: US-09-668-419-1 from: 1 to: 267

```

1  TACACAGATGTCGAGCGAGTACCAGCGTAT ... AGGACACAGATCCGCT 47
123  TysrAsprgIsuencIuolIneunlaetgIsueneuarIguIsanIlegIule 139
48  GGGATATAGAGGAACATCATTCGATCCGCTCGGGCGCTTCGTTAT ..... 90
139  uGIyAsngIyProIeouGIuIaIlaIleSerAlaIeunIyTyrIySerI 156
91  .. CCAGGCGGACAGCACCCGGGGCCCAAGCGTCGTCCCTATTAATCCGAT 138
156  hrtGIyGIyThrtGIuIeuprOthrtleuAlaIargrSerPhelIleIcYsIle 172
139  CAGATGATCTCCGAGCGCCGCGAGATTCATCCATCCTTTTGAGGGCTCG 188
173  GlmetIleSerGIuIaIaIarphheGInIyrtIleGIuGIyGImetAr 189
189  CCAATACATTAACAGCGGGGAGTCATTTCTTCGCCAGATACATACGTGCG 238
189  gYthrtArGIleArGIYrAshtArGrSerAlaIarOspProSerValIleT 206
239  AGCTGAGACATAGTGGGGCCACAACATCCAGCAGATCCAGACTTCACG 288
206  hrtIeugIuIsanSerItpGJIyArGIleuSerThrtAlaIleGIuIuSerAsn 222
289  GATGCGCGTPTTAAATTAACCATTCGTGGTGGGTATATCCACCGCTACTT 338
223  GInGIyAlaIarheAlaSerIroIleGIuIeunGIaIarGrAsnGIySerLy 239
339  CGTGAAGTCAGACAAATGTTCCGACAGTGAATGCACAGCTTACGATCAATG 388
239  sPhSerValIyTyrAspValSerIleIeunIleproIleIleAlaIeunetv 256
389  TGTTGTATGTAGAGGACCGACCATCTTCC 417
256  aIIytrArGysAlaIarProIProSer 265

```



```

1  TACACGATCTGGAGCGGATACGCCGGTGTATATGGAGCCAGATCCCTCTGGG  50
   ||| |||||
115  TYPProAspLeuGlnIaArgTyrIaGlnIaHisArgAspGlnIleProLeuG  131
   ||| :|||
131  YIleAspGlnIleuIleGlnIleSerValThiAlaIleuArgPheProGly  148
   ||| :|||
148  GCACCCGGGGCCCAAGCTCGTTCCTTATATCTCATTCAGATGATCTCC  150
   ||| |||||
148  eRthArgThiArgIaIaArgSerIleuIleuIleGlnIleMetIleSer  164
   ||| |||||
151  GAGCGCCGCGAGATTCAATCCCATCTTTTGGAGGCGCTGCCCAATCAATTA  200
   ||| |||||
165  GlnIaIaIaArgPheAsnProIleuTyrArgTyrIaGlnTyrIleAs  181
   ||||| |||||
181  nSerGlyAlaSerPheIleuProAspValTyrMetLeuIuIeuGlnIhR  198
   ||||| |||||
201  GTTGGGGCCCAACAATCCAGCAGATTCACAGCTTACCGATGGCGTTT  300
   ||||| |||||
198  eRTrpGlyGlnIleSerThiGlnIaGlnIleSerThiAspGlyValPhe  214
   ||||| |||||
251  GATTAACCCATTGGGTGGGTATATGACCAGGCTAACTCGAGAGTTGAG  350
   ||||| |||||
215  AsnAsnProIleArgIleuAlaIleProProGlyAsnPheValThiLeuTh  233
   ||||| |||||
351  CAATGTTGCGAGGTGATCGCAGCTTACCGATCAATGTTGTTGTATGTA  400
   ||||| |||||
231  rAsnValAlaArgAspAlaIleAlaSerIleuAlaIleMetIleuPheVal  246
   ||||| |||||
401  GGGAGCCGCAATCTTCTCC  420
   ||||| |||||
248  GlyGlnArgProSerSerSer  254

```

```
seq_name: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:US-09-601-667B-39
seq_documentation_block:
; Sequence 39, Application US/09601667B
; GENERAL INFORMATION:
; APPLICANT: Morris, Peter
; APPLICANT: Stiefel, Thomas
; APPLICANT: Voelter, Wolfgang
; APPLICANT: Welters, Peter
; TITLE OF INVENTION: Recombinant Mistletoe Lectins
; FILE REFERENCE: 29841/36636
; CURRENT APPLICATION NUMBER: US/09/601,667B
; PRIOR APPLICATION NUMBER: PCT/EP99/00696
; PRIOR FILING DATE: 1999-02-03
; PRIOR APPLICATION NUMBER: D 198 04 210.8
; PRIOR FILING DATE: 1998-02-03
; NUMBER OF SEQ ID NOS: 41
; SEQ ID NO 39
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mistletoe lectin A1 (matched)
; US-09-601-667B-39

alignment_scores:
  Quality: 604.50      Length: 140
  Ratio: 4.545         Gaps: 1
  Percent Similarity: 95.000  Percent Identity: 85.000

alignment_block:
US-09-627-165B-15 x US-09-601-667B-39  ..

Align seg 1/1 to: US-09-601-667B-39 from: 1 to: 253

1  TACACAGATCTGAGGATACGCGGTCATAGGACCAATCCCTGTGGG 50
115 TyrProAspLeuGlnAlaArgSerIleLeuIleGlnMetIleSer 130
51  TATAGAGCAATCATTCATCGCTCGGCGCTCTATACGAGCGGCA 100
130 TTTAAGAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 147
101  GCACCGCGGCCCAAGCTCGTTCCTTATATCTCATTCAGATGATCTCC 150
147 errTATGTGTTGlnAlaArgSerIleLeuIleGlnMetIleSer 163
151  GAGGCGCGGATTCATCCATCTTTGGAGGCGTCCGCAATACATTA 200
164 GlnAlaAlaArgPheAsnProIleLeuTyrArgGlnTyrIleAs 180
201  CACGCGGAGCTCATTTCTCCGACATGTACATGCTCGAGCTGAGACTA 250
180 nSerGlnAlaSerPheLeuProAspValTyrMetLeuGlnLeuI 197
251  GTTGGGGCCCAATCCAGCAAGTCCAGCTACGATGCGCTTTT 300
197 errTPTGlnGlnSerThrGlnValGlnHisSerThrAspIleVal 213
301  AATTAACCATTTGCGTGGGTATATCCACCGGTAATCTTGTGAGCTGAG 350
214 AsnAsnProIleArgLeuAlaIleProProGlnAsnPheValThrLeu 230
351  CAATGTTCGCGAGCTGATCCAGCTTACGCTTACGCTGTTGTGTATGA 400
230 TAsnValAlaArgAspValIleAlaSerLeuAlaIleMetLeuPheValCysG 247
401  GGGACCGACCATCTTCCTCC 420
247  TGTGlnArgProSerSerSer 253
```

```
seq_name: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:US-09-601-667B-2
seq_documentation_block:
; Sequence 2, Application US/09601667B
; GENERAL INFORMATION:
; APPLICANT: Morris, Peter
; APPLICANT: Stiefel, Thomas
; APPLICANT: Voelter, Wolfgang
; APPLICANT: Welters, Peter
; TITLE OF INVENTION: Recombinant Mistletoe Lectins
; FILE REFERENCE: 29841/36636
; CURRENT APPLICATION NUMBER: US/09/601,667B
; PRIOR APPLICATION NUMBER: PCT/EP99/00696
; PRIOR FILING DATE: 1999-02-03
; PRIOR APPLICATION NUMBER: D 198 04 210.8
; PRIOR FILING DATE: 1998-02-03
; NUMBER OF SEQ ID NOS: 41
; SEQ ID NO 2
; LENGTH: 255
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: M1A-chain
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 15
; OTHER INFORMATION: product= "Xaa is Asp or Glu"
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 63
; OTHER INFORMATION: product= "Xaa is Gly or Gln"
; OTHER INFORMATION: /label= Xaa2
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 66
; OTHER INFORMATION: product= "Xaa is Ile or Val"
; OTHER INFORMATION: /label= Xaa3
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 75
; OTHER INFORMATION: product= "Xaa is Leu or Ala"
; OTHER INFORMATION: /label= Xaa4
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 107
; OTHER INFORMATION: product= "Xaa is missing"
; OTHER INFORMATION: /label= Xaa5
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 113
; OTHER INFORMATION: product= "Xaa is Asn or Thr"
; OTHER INFORMATION: /label= Xaa6
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 117
; OTHER INFORMATION: product= "Xaa is Pro or Thr"
; OTHER INFORMATION: /label= Xaa7
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 141
; OTHER INFORMATION: product= "Xaa is Asp or Glu"
; OTHER INFORMATION: /label= Xaa8
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 141
; OTHER INFORMATION: product= "Xaa is Ser or Thr"
; OTHER INFORMATION: /label= Xaa9
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 145
```



```
OTHER INFORMATION: product= "Xaa is Phe or Tyr"
OTHER INFORMATION: /label= Xaa10
FEATURE:
NAME/KEY: SITE
LOCATION: 152
OTHER INFORMATION: product= "Xaa is Thr or Ala"
OTHER INFORMATION: /label= Xaa11
FEATURE:
NAME/KEY: SITE
LOCATION: 177
OTHER INFORMATION: product= "Xaa is Ala or Tyr"
OTHER INFORMATION: /label= Xaa12
FEATURE:
NAME/KEY: SITE
LOCATION: 180
OTHER INFORMATION: product= "Xaa is Tyr or Asp"
OTHER INFORMATION: /label= Xaa13
FEATURE:
NAME/KEY: SITE
LOCATION: 185
OTHER INFORMATION: product= "Xaa is Ala or Glu"
OTHER INFORMATION: /label= Xaa14
FEATURE:
NAME/KEY: SITE
LOCATION: 191
OTHER INFORMATION: product= "Xaa is Val or Met"
OTHER INFORMATION: /label= Xaa15
FEATURE:
NAME/KEY: SITE
LOCATION: 219
OTHER INFORMATION: product= "Xaa is Ile or Phe"
OTHER INFORMATION: /label= Xaa16
FEATURE:
NAME/KEY: SITE
LOCATION: 224
OTHER INFORMATION: product= "Xaa is Pro or Ser"
OTHER INFORMATION: /label= Xaa17
FEATURE:
NAME/KEY: SITE
LOCATION: 225
OTHER INFORMATION: product= "Xaa is Pro or Thr"
OTHER INFORMATION: /label= Xaa18
FEATURE:
NAME/KEY: SITE
LOCATION: 232
OTHER INFORMATION: product= "Xaa is Thr or Ser"
OTHER INFORMATION: /label= Xaa19
FEATURE:
NAME/KEY: SITE
LOCATION: 236
OTHER INFORMATION: product= "Xaa is Asp or Ser"
OTHER INFORMATION: /label= Xaa20
US-09-601-667B-2

alignment_scores:
Quality: 599.00      Length: 140
Ratio: 4.538         Gaps: 0
Percent Similarity: 94.286      Percent Identity: 84.286

alignment_block:
US-09-627-165B-15 x US-09-601-667B-2 ..
Align seg 1/1 to: US-09-601-667B-2 from: 1 to: 255

1 TACACAGATCTGGAGGATACGCCGCTCATAGGACAGATCCCTCTGGG 50
116 Tyr***AspleuGluArgTyrAlaGlyHisArgAspGlnIleProLeuG1 132
51 TATAGAGCACTCATTCATTCGCTCGCGGCTTCCTTATCCAGGGCGCA 100
132 ylle**GlnleuIleGlnserVal**AlaleuArg**ProGlyGlys 149
```

```
101 GACCCGGCCAGCTCGTCCCTTAATACCTCATTCAGATGATCTCC 150
149 erThrArg***GlnAlaArgSerIleleuIleGlnMetIleSer 165
151 GAGCCCGGAGATTCAATCCATCTTTTGGAGGCGTCGCCAATACATTA 200
166 GluAlaAlaArgPheAsnProIleleuIleTrpArg***ArgGln***IleAs 182
201 CAGCCGGGAGTCATCTTCGCCAGATGACATGCTCAGCTGAGAGCTA 250
182 nserGly***SerPheLeuProAsp***TyrMetLeuGluLeuGluThrS 199
251 GTTGGGCCCAACATCCAGCAGATGTCAGCTGAGAGGCGCTTTT 300
199 erTrpGlyGlnGlnSerThrGlnVal3InHisSerThrAspGlyValPhe 215
301 AATRACCATTTCCGTTGGTATATCCACCGGTAACTTCGTGAGCTTGA 350
216 AsnAsnPro***ArgLeuAlaIle****GlyAsnPheValThrLeu** 232
351 CAATGTCGCGAGCTGATCGCCAGCTTACGATCATGTTGTTGTATGTA 400
232 *AsnValArg***ValIleAlaSerLeuAlaIleMetLeuPheValCysG 249
401 GGAACCGACCATCTCTCTCC 420
249 yGluAlaArgProSerSerSer 255

seq_name: /cgn2_6/prodata/1/paa/US09_NEW_COMB.pep:US-09-601-667B-41

seq_documentation_block:
; Sequence 41, Application US/0960167B
; GENERAL INFORMATION:
; APPLICANT: Morris, Peter
; APPLICANT: Stiefel, Thomas
; APPLICANT: Voelters, Peter
; TITLE OF INVENTION: Recombinant Mistletoe Lectins
; FILE REFERENCE: 29841/36636
; CURRENT FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: PCT/EP99/00696
; PRIOR FILING DATE: 1999-02-03
; PRIOR APPLICATION NUMBER: D 198 04 210.8
; PRIOR FILING DATE: 1998-02-03
; NUMBER OF SEQ ID NOS: 41
; SEQ ID NO 41
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: MUA-chain
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (15)..(15)
; OTHER INFORMATION: Xaa is Asp or Glu
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (66)..(66)
; OTHER INFORMATION: Xaa is Ile or Val
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (75)..(75)
; OTHER INFORMATION: Xaa is Leu or Ala
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (114)..(114)
; OTHER INFORMATION: Xaa is Asn or Thr
; FEATURE:
```

```

NAME/KEY: SITE
LOCATION: (118)..(118)
OTHER INFORMATION: xaa is Pro or Thr
FEATURE:
NAME/KEY: SITE
LOCATION: (135)..(135)
OTHER INFORMATION: xaa is Asp or Glu
FEATURE:
NAME/KEY: SITE
LOCATION: (142)..(142)
OTHER INFORMATION: xaa is Ser or Thr
FEATURE:
NAME/KEY: SITE
LOCATION: (146)..(146)
OTHER INFORMATION: xaa is Phe or Tyr
FEATURE:
NAME/KEY: SITE
LOCATION: (153)..(153)
OTHER INFORMATION: xaa is Thr or Ala
FEATURE:
NAME/KEY: SITE
LOCATION: (178)..(178)
OTHER INFORMATION: xaa is Ala or Tyr
FEATURE:
NAME/KEY: SITE
LOCATION: (181)..(181)
OTHER INFORMATION: xaa is Tyr or Asp
FEATURE:
NAME/KEY: SITE
LOCATION: (186)..(186)
OTHER INFORMATION: xaa is Ala or Glu
FEATURE:
NAME/KEY: SITE
LOCATION: (192)..(192)
OTHER INFORMATION: xaa is Val or Met
FEATURE:
NAME/KEY: SITE
LOCATION: (220)..(220)
OTHER INFORMATION: xaa is Ile or Phe
FEATURE:
NAME/KEY: SITE
LOCATION: (225)..(225)
OTHER INFORMATION: xaa is Pro or Ser
FEATURE:
NAME/KEY: SITE
LOCATION: (226)..(226)
OTHER INFORMATION: xaa is Pro or Thr
FEATURE:
NAME/KEY: SITE
LOCATION: (233)..(233)
OTHER INFORMATION: xaa is Thr or Ser
FEATURE:
NAME/KEY: SITE
LOCATION: (237)..(237)
OTHER INFORMATION: xaa is Asp or Ser
US-09-601-667B-41

```

```

alignment_scores:
Quality: 599.00      Length: 140
Ratio: 4.538        Gaps: 0
Percent Similarity: 94.286      Percent Identity: 84.286

```

alignment_block:

US-09-627-165B-15 x US-09-601-667B-41 ..

Align seg 1/1 to: US-09-601-667B-41 from: 1 to: 256

```

1 TACACGATCTGAGAGATACGCGGTATAGGAGACGATCCCTGAGG 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
117 Tyr***AspLeuGluArgTyrAlaGlyHisArgAspGlnIleProLeuG 133
51 TATAGAGAGACTCATTCATCCGCTCGGCGCTTCGTATCCAGGCGGCA 100

```

```

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
133 Ylle***GlnIleuIleGlnSerVal***AlaLeuArg***ProGlyGlyS 150
101 GCACCCGGGCCCCAAGCTCGTTCCTTATATATCTCATTCAGATGATCTCC 150
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
150 eXThrArg***GlnAlaArgSerIleIleuIleGlnIleSer 166
151 GAGCCGGAGATTCATCCATCTTTGAGGGCTCCCAATACATTA 200
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
167 GluAlaIleArgPheAsnProIleuLeuTyrParg***ArgGln***IleAs 183
201 CAGCGGGAGTCATTTCTCCGACATGACATGCTGACGCTGAGACTA 250
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
183 nserGly***SerPheLeuProAsp***TyrIleLeuGlnLeuGlnTyrS 200
251 GTTGGGCCCAACATTCAGCGAAGTCCAGAGCTTACGATGGCGTTTGT 300
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
200 eXTrpGlyGlnGlnSerThrGlnValGlnHisSerThrAspGlyValPhe 216
301 AATAACCCATTTGCGTGGGATATCCACGGGTAACCTTCGTGACGTTGAG 350
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
217 AsnAsnPro***ArgLeuAlaIle****GlyAsnPheValThrIleu*** 233
351 CATGTTCGCGACGTGATCCGACGCTTACGATCATGTTGTTGATGTA 400
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
233 *AsnValArg***ValIleAlaSerLeuAlaIleMetLeuPheValGlySg 250
401 GCGACGACCATCTTCTCC 420
:::|||||:|||||:|||||
250 YgluArgProSerSerSer 256

```

seq_name: /cgn2_6/prodata/1/paa/US10_NEW_COMB.pep:US-10-083-336A-1

```

seq_documentation_block:
Sequence 1, Application US/10083336A
GENERAL INFORMATION:
APPLICANT: Olson, Mark A
APPLICANT: Millard, Charles B
APPLICANT: Byrne, Michael P
APPLICANT: Wannemacher, Robert W
TITLE OF INVENTION: Rizin Vaccine and Methods of Making and Using Thereof
FILE REFERENCE: P67452050 (RIID 01-58)
CURRENT APPLICATION NUMBER: US/10/083,336A
CURRENT FILING DATE: 2002-05-21
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 576
TYPE: PRT
ORGANISM: Rictinus communis
US-10-083-336A-1

```

```

alignment_scores:
Quality: 259.50      Length: 178
Ratio: 2.257        Gaps: 5
Percent Similarity: 64.607      Percent Identity: 41.011

```

alignment_block:

US-09-627-165B-15 x US-10-083-336A-1 ..

Align seg 1/1 to: US-10-083-336A-1 from: 1 to: 576

```

1 TACACGATCTGAGAGATACGCGGTATAGGAGACGATCCCTGAGG 47
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
158 TyrAspArgLeuGlnGlnIleuAlaGlyAsnLeuArgGlnAsnIleGluLe 174
48 GGGTATAGAGAGACTCATTCATCCGCTCGGCGCTTCGTAT ..... 90
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
174 uGlyAsnGlyProLeuGlnGlnAlaIleSerAlaLeuTyrTyrIleSer 191
91 ..CCAGGCGGACGACCCGGGCCCAAGCTCGTTCCTTATATCTCATTT 138
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

```

191 hrclyglythrGlnLeuProThrLeuAlaArgSerPheIleIleCysIle 207
139 CAGATGATCTCCGAGCCGCGAGATTCATCCATCTTTTGAGGCGTCG 188
|||||
208 GlnMetIleSerGluAlaAlaArgPheGlnIleTyrIleGluGlyIleMetArg 224
189 CCAATACATTACAGCGGGAGTCATTTCTTCCGACATGATGATGCTCG 238
|||||
224 gthrrArgIleArgTyrAsnArgArgSerAlaProAspProSerValIleT 241
239 AGCTGAGACTAGTTGGGGCCACAAATCCAGCAATGCGAGCTTAGC 288
|||||
241 hrlengGluAsnSerTrpGlyArgLeuSerThrAlaIleGlnGlnSerAsn 257
289 GATGGCGTTTAAATACCACTTCGTTGGGTGATTCACACCGTAACCT 338
|||||
258 GlnGlyAlaPheAlaSerProIleGlnLeuGlnArgArgAsnGlySerIly 274
339 CGTACGCTTGACCAATGTTCGCGAGCTGATCGCAGCTTAGCATCAT 388
|||||
274 sPheSerValIlyrAspValSerIleLeuIleProIleIleAlaLeuMetV 291
389 TGTTGTATGTAGAGGACCGACCATCTCTCCGAGCTCGCTATTGGCG 438
|||||
291 alTyrArgCysAlaProProProSerSerGln.....PheSer 303
439 CTGGTCTATGACCGCTTTCGAAATAATAGCGCGCGCTGACGATGTAC 488
|||||
304 LeuLeuIleArgProValIleAlaProAsnPheAsnAla.....AspVal.. 317
489 CTGCATGCTTCGAAACCCACCGCTGCGCATCGTA 522
|||||
318 .CysMetAspProGlnProIleValArgIleVal 328

seq_name: /cgn2_6/prodata/1/paa/US10_NEW_COMB.pep:US-10-127-890-1

seq_documentation_block:
; Sequence 1, Application US/10127890
; GENERAL INFORMATION:
; APPLICANT: Bectel, Marc D.
; Studnka, Gary M.
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
; Proteins
; NUMBER OF SEQUENCES: 173
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Mcandrews, Held & Malloy, Ltd.
; STREET: 500 West Madison Street, 34th floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60661
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/127,890
; FILING DATE: 23-Apr-2002
; CLASSIFICATION: <unknown>
; PRIORITY DATA:
; APPLICATION NUMBER: US/08/646,360
; FILING DATE: 13-MAY-1996
; APPLICATION NUMBER: PCT/US94/05348
; FILING DATE: 12-MAY-1994
; APPLICATION NUMBER: US 08/064,691
; FILING DATE: 12-MAY-1993
; APPLICATION NUMBER: US 07/988,430
; FILING DATE: 09-DEC-1992
; APPLICATION NUMBER: US 07/901,707
; FILING DATE: 19-JUN-1992
; APPLICATION NUMBER: US 07/787,567

```

```

; FILING DATE: 04-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McNicholas, Janet M.
; REGISTRATION NUMBER: 32,918
; REFERENCE/DOCKET NUMBER: 200-70..P4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/707-9155
; TELEFAX: 312/707-9155
; TELEX: 650 388-1248
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 267 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-127-890-1

alignment_scores:
Quality: 217.00 Length: 143
Ratio: 2.309 Gaps: 2
Percent Similarity: 65.734 Percent Identity: 39.860

alignment_block:
US-09-627-165B-15 x US-10-127-890-1 ..

Align seg 1/1 to: US-10-127-890-1 from: 1 to: 267

1 TACACAGATCTGAGCGATACGCCGTCAT...AGGACCAATCCCTCT 47
|||||
123 TyrAspArgLeuGlnGlnLeuAlaGlyAsnLeuArgGlnAsnIleGluLe 139
48 GGGTATAGAGCACTCAATCCCTCTCGCGCTCAT... 90
|||||
139 uGlyAsnGlyProLeuGlnGlnAlaIleSerAlaLeuTyrTyrSerT 156
139 ..CCAGCGCGACACCGCGGCCAGCTCGTCCCTTAATTCATCATT 138
|||||
156 hrclyglythrGlnLeuProThrLeuAlaArgSerPheIleIleCysIle 172
139 CAGATGATCTCCGAGCGCGCGATTCATCCATCTTTTGAGGCGTCG 188
|||||
173 GlnMetIleSerGluAlaAlaArgPheGlnIleTyrIleGluGlyIleMetArg 189
189 CCAATACATTACAGCGGGAGTCATTTCTTCCGACATGATGATGCTCG 238
|||||
189 gthrrArgIleArgTyrAsnArgArgSerAlaProAspProSerValIleT 206
239 AGCTGAGACTAGTTGGGGCCAAATCCAGCAATGCGAGCTTAGC 288
|||||
206 hrlengGluAsnSerTrpGlyArgLeuSerThrAlaIleGlnGlnSerAsn 222
289 GATGGCGTTTAAATACCACTTCGTTGGGTGATTCACACCGTAACCT 338
|||||
223 GlnGlyAlaPheAlaSerProIleGlnLeuGlnArgArgAsnGlySerIly 239
339 CGTACGCTTGACCAATGTTCGCGAGCTGATCGCAGCTTAGCATCAT 388
|||||
239 sPheSerValIlyrAspValSerIleLeuIleProIleIleAlaLeuMetV 256
389 TGTTGTATGTAGGAGCGACCATCTTCC 417
|||||
256 alTyrArgCysAlaProProProSerSer 265

seq_name: /cgn2_6/prodata/1/paa/US09_NEW_COMB.pep:US-09-726-651A-4

seq_documentation_block:
; Sequence 4, Application US/0972651A
; GENERAL INFORMATION:
; APPLICANT: Danson, Jon
; Applicant: DANSON, WILLIAM O.
; APPLICANT: GRANTHAM, GEORGE L.

```


alignment_block:
US-09-627-165b-15 x US-10-127-890-7 ..

Align seg 1/1 to: US-10-127-890-7 from: 1 to: 263

```

1 TACACAGATCTGAGCGATACGCCGTCAT...AGGACACAGATCCCTCT 47
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
111 TyrgluarGluGlnIleAlaIaIaGlyLysProArGlnIlysIleProIle 127
48 GGGTTATAGAGCAATTCATCCGTCCTCGCGGCTTCGTTATCCAGGCG 97
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
127 eGlyLeuProAlaIeuaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 142
98 GCAGCAGCCGGGCGCCAGCTCTGCTCCCTATATCCCTCATTCAGATGATC 147
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
142 yAspSerThrAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 158
148 TCCGAGCCGGCGAGATTCATCCCTTTTGAGGGCTCGCCATATACAT 197
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
159 AlaGlnAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 175
198 TAACAGCCGGGAGTCATTTCTTCCGACATGATACGTCGAGGTGAGA 247
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
175 aTyArGAspGlu.....ValProSerLeuAlaThrIleSerLeuGln 190
248 CTAGTTGGGGCCCAACATCCAGCAAGTCCAG.....CAGCTTACGGAT 291
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
190 snSerTIPSerGlyLeuSerLysGlnIleGlnLeuAlaIaIaIaIaIaIa 206
292 GCGGTTTTTAATACCATTTGCGTTGGGTATATCCACCGGTAACTTGT 341
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
207 GlyIlePheArGthrProIleValIaIaIaIaIaIaIaIaIaIaIaIaIa 223
342 GACGTTAGCAAGTTCGCGAGCTGATCCGACCTTAGCATATGTTGT 391
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
223 GlnIleThrAsnValIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 240
392 TTGTA 396
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
240 euIeu 241

```

seq_name: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep:US-10-127-890-6

```

seq_documentation_block:
; Sequence 6, Application US/10127890
; GENERAL INFORMATION:
; APPLICANT: Better, Marc D.
;               Carroll, Stephen F.
;               Studnika, Gary M.
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
;                   Proteins
; NUMBER OF SEQUENCES: 173
; CORRESPONDENCE ADDRESS:
; ADDRESS: McAndrew, Held & Malloy, Ltd.
; STREET: 500 West Madison Street, 34th floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60661
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/127.890
; FILING DATE: 23-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION NUMBER: US/08/646,360
; APPLICATION NUMBER: PCT/US94/05348
; FILING DATE: 12-MAY-1994

```

```

; APPLICATION NUMBER: US 08/064,691
; FILING DATE: 12-MAY-1993
; APPLICATION NUMBER: US 07/988,430
; FILING DATE: 09-DEC-1992
; APPLICATION NUMBER: US 07/901,707
; FILING DATE: 19-JUN-1992
; APPLICATION NUMBER: US 07/787,567
; FILING DATE: 04-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: MONICHOIAS, Janet M.
; REGISTRATION NUMBER: 32,918
; REFERENCE/DOCKET NUMBER: 200-70.P4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/707-9155
; TELEFAX: 312/707-9155
; TELEX: 650 388-1248
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 247 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-127-890-6

```

alignment_scores:
Quality: 153.50 Length: 135
Ratio: 1.651 Gaps: 5
Percent Similarity: 68.889 Percent Identity: 32.593

alignment_block:
US-09-627-165b-15 x US-10-127-890-6 ..

Align seg 1/1 to: US-10-127-890-6 from: 1 to: 247

```

1 TACACAGATCTGAGCGATACGCCGTCAT...AGGACACAGATCCCTCT 47
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
111 TyrgluarGluGlnIleAlaIaIaGlyLysIleArGlnIlysIleProIle 127
48 GGGTTATAGAGCAATTCATCCGTCCTCGCGGCTTCGTTATCCAGCG 97
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
127 uGlyLeuProAlaIeuaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 144
98 GCAGCAGCCGGGCGCCAGCTCGTCCCTTATATCCCTCATTCAGATGATC 147
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
144 IaaSerAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 158
148 TCCGAGCCCGGAGATTCATCCCATCTTTGAGGGCTCGCCATATACAT 197
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
159 SerGlnAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 175
198 TAACAGCCGGGAGTCATTTCTTCCGACATGATACGTCGAGGTGAGA 247
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
175 IAsp.....LysThrPheLeuProSerLeuAlaIaIaIaIaIaIaIaIa 190
248 CTAGTTGGGGCCCAACATCCAGCAAGTCCAG.....CAGCTTACGGAT 291
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
190 snSerTIPSerAlaLeuSerLysGlnIleGlnIleAlaIaIaIaIaIaIa 206
292 GCGGTTTTTAATACCATTTGCGTTGGGTATATCCACCGGTAACTTGT 341
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
207 GlyIlePheGlnIleuSerProIaIaIaIaIaIaIaIaIaIaIaIaIaIa 223
342 GACGTTAGCAATGTTCCGAC.....GTATAGCCAGCTGATGATCA 385
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
223 aThrIleThrAsnValaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 240
386 TGTTG 390
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
240 euIeu 241

```

seq_name: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep:US-10-127-890-9

seq_documentation_block:
; Sequence 9, Application US/10127890
; GENERAL INFORMATION:
; APPLICANT: Better, Marc D.
; Carroll, Stephen F.
; Studnika, Gary M.
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
; Proteins
; NUMBER OF SEQUENCES: 173
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.
; STREET: 500 West Madison Street, 34th floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60661
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/127,890
; FILING DATE: 23-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,360
; FILING DATE: 13-MAY-1996
; APPLICATION NUMBER: PCT/US94/05348
; FILING DATE: 12-MAY-1994
; APPLICATION NUMBER: US 08/064,691
; FILING DATE: 12-MAY-1993
; APPLICATION NUMBER: US 07/988,430
; FILING DATE: 09-DEC-1992
; APPLICATION NUMBER: US 07/901,707
; FILING DATE: 19-JUN-1992
; APPLICATION NUMBER: US 07/787,567
; FILING DATE: 04-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McNicholas, Janet M.
; REGISTRATION NUMBER: 32,918
; REFERENCE/DOCKET NUMBER: 200-70.P4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/707-8889
; TELEFAX: 312/707-9155
; TELEX: 650 368-1248
; COUNTRY: USA
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 261 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-10-127-890-9

alignment_scores:
Quality: 152.50 Length: 135
Ratio: 1.694 Gaps: 3
Percent Similarity: 66.667 Percent Identity: 30.370

alignment_block:
US-09-627-165b-15 x US-10-127-890-9 ..

Align seg 1/1 to: US-10-127-890-9 from: 1 to: 261

1 TACACGATCTGTGAGGATACCGCGT.....CATGAGGACGATCCC 44
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
122 TyrProthrleuclulysylsalaclyvalphSerArgasnclvalcl 138
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
45 TCRTGGATGAGGAGACTGATTCGCGTCTGGCGCCCTTCGTTATCCAG 94
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

138 nLeuGlylleGlnlleuSerSeraspIleGlyllysIleSerclGln 155
95 GCGGCGAGCACC...CGGCGCCAGCTCGTTCCCTTATATCCCTCATTCAG 141
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
155 lYserPheThrclulyslleGlnAlaaspPheleuValAlaIleGln 171
142 ATGATCTCCGAGCGCGGAGATTCATCCATCCATCTTTGGAGGCGCCCA 191
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
172 MetvalSerGlnAlaAlaArgPheylsTYlleGlnuAsnGlnValys.. 187
192 ATACATTAACACCGGCGGAGTCATCTCTCCGACATGACATGCTCGAGC 241
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
188 ...ThrasnPhenasnArgaspPheSerProksnAspIysValuAspL 203
242 TGGAGACTAGTTGGGCGCAACATCCAGCAGATCCAGCAGCTACGAT 291
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
203 euGlnclulasntrpGlyllysIleSerThrAlaIleHisasnSerlyasn 219
292 GCGGCTTTTAAATACCCATTTGCGGTGGATATCCACCGGACTTCGCT 341
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
220 GlyAlaIeuProIysProIeuGlnleuIysasnAlaaspGlyThrlySTR 236
342 GACGTTGAGCAATGTTCCGCGACGCGATCCGACGCTTACGATCATGTTGT 391
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
236 PilevalleuArgvalaspIulleyspProaspvalGlyleuLeuasnT 253
392 TTGCTA 396
||| |||||
253 yval 254

seq_name: /cgn2_6/prodata/1/paa/US10_NEW_COMB.pep:US-10-127-890-102

seq_documentation_block:
; Sequence 102, Application US/10127890
; GENERAL INFORMATION:
; APPLICANT: Better, Marc D.
; Carroll, Stephen F.
; Studnika, Gary M.
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
; Proteins
; NUMBER OF SEQUENCES: 173
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.
; STREET: 500 West Madison Street, 34th floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60661
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/127,890
; FILING DATE: 23-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,360
; FILING DATE: 13-MAY-1996
; APPLICATION NUMBER: PCT/US94/05348
; FILING DATE: 12-MAY-1994
; APPLICATION NUMBER: US 08/064,691
; FILING DATE: 12-MAY-1993
; APPLICATION NUMBER: US 07/988,430
; FILING DATE: 09-DEC-1992
; APPLICATION NUMBER: US 07/901,707
; FILING DATE: 19-JUN-1992
; APPLICATION NUMBER: US 07/787,567
; FILING DATE: 04-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McNicholas, Janet M.
; REGISTRATION NUMBER: 32,918

Page 9

20T-069-17T-0T-50

alignment_block:

Align seg 1/1 to: US-10-127-890-102 from: 1 to: 251

```

1  TACACAGATCTGGAGCGATACGCCGCTCATAGGAGACCAATCCCTGGG 50
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
113  TYPProSerLeuGluGluGlyGluValysAlaTyrArgGluThrAspLeuCl 129
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
51  TATPAGGAGAACGC.....ATTCAATCCGCTCGGGCGTTCGTATC 91
   ||| ||||| ||| ||||| :||| :||| :||| :||| :||| :|||
129  YLleGluProLeuArgIleGlyIleTyrSlyLeuAspGluAsnAlaIleA 146
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
92  CAGCGGGAGACACCGGGCCCACTCGTCCCTTAATACCTCAATTCAG 141
   :: :: :: ||| ||||| :||| :||| :||| :||| :||| :|||
146  sPaSnArYrAspProThrGluIleAspSerLeuValIValIleGln 162
   :: :: :: ||| ||||| :||| :||| :||| :||| :||| :|||
142  ATATATCTCCGAGCGCGCAGATTCAATCCATCTTTTGGAGGCGTCCCA 191
   :: :: :: ||| :||| :||| :||| :||| :||| :|||
163  MeValSerGluAlaIleAlaIleArgPheThrPheIleGluAsnGlnIleArgAs 179
   :: :: :: ||| :||| :||| :||| :||| :||| :|||
192  ATACATTAAACAGCGGGAGTACTATTCTTCCTCCGACATGCATCATCGTCGAC 241
   :: :: :: ||| :||| :||| :||| :||| :||| :|||
179  n.....AspPheGlnGluArgIleArgProIleAsnAsnThrIleSerL 194
   :: :: :: ||| :||| :||| :||| :||| :||| :|||
242  TGCAGACTACTTGGCGCCCAACATCCACGCCAGTCCAGCACTCT..ACG 288
   :: :: :: ||| :||| :||| :||| :||| :||| :|||
194  euGluAsnLysTrpGlyLysLeuSerPheGlnIleArgThrSerGlyVala 210
   :: :: :: ||| :||| :||| :||| :||| :||| :|||
289  GATGGCGCTTTTAATACCATTCGTTGGGTGGGTATATCCACCGGTAACTT 338
   ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
211  AsnGlnMetPheSerTrpGluAlaValGluLeuGluValArgAlaAsnGlyLysLy 227
   ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
339  CGTGCAGCTTGGACCAATGTTACGCGAGCTGATCCGACGTTAGAGATCAAGT 388
   ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
227  sTyTrpValThrAlaValAspGlnValLysProLysIleAlaLeuLeuL 244
   ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
389  TGTTTGATCTGACGAGC 405
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
244  ySPheValCysLysAsp 249
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

seq_name: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep:US-10-127-890-111

seq_documentation_block:
 sequence: 111 Application NS/10127890

GENERAL INFORMATION:

APPLICANT: Better, Marc D.
Carroll, Stephen F.

Studnıka, Gary M.

TITLE OF INVENTION: IMMUNOPROTECTANT

NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA

APPLICATION NUMBER: 08/06/0-07-
FILING DATE: 13-MAY-1996

APPLICATION NUMBER: PCT/US94/0534
FILING DATE: 12-MAY-1994

APPLICATION NUMBER: US 08/064,691
FILING DATE: 12-MAY-1993

FILING DATE: 12 MAR 1999
 APPLICATION NUMBER: US 07/988,430

FILING DATE: 09-DEC-1992
APPLICATION NUMBER: US 07/901,701

FILING DATE: 19-JUN-1992
APPLICATION NUMBER: US 07/787,56

APPLICATION NUMBER: 04-11991
FILING DATE: 04-NOV-1991

ATTORNEY/AGENT INFORMATION:
NAME: McNicholas, Janet M.

REGISTRATION NUMBER: 32,918
REFERENCE/DOCKET NUMBER: 200-70.

TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/707-8489

TELEPHONE: 312/707-9999
TELEFAX: 312/707-9155

TELEX: 650 388-1248
INFORMATION FOR SEQ ID NO: 111:

SEQUENCE CHARACTERISTICS:
LENGTH: 251 amino acids

LENGTH: 201 amino acids
TYPE: amino acid

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 111
87-890-111

2
/
C
C
4
4

ent_scores:

Quality:	147.00	Length
Ratio:	1.690	Ga

at Similarity: 62.590 Percent Identity

ent_block: 10 107 000-111

627-165B-15 x US-10-12/-890-111 ..

seg 1/1 to: US-10-127-890-111 from:

1 TACACAGATCTGGAGCGATACGCCGTCATAGGAC

113 TyrProSerLeuGluGlyGluLysAlaTyrArgGlu

41 TATAGAGGACTC.....ATTCATCCGTC

[illegible]

29 yilegluproleuarglrelyrelyscyscann

92 CAGGCGGACGACCCGGGCCAAGCTCGTTCCTT

46 spasnTyrlYsProThrgLlLeaLaSerSerLeu

42 ATGATCTCCGAGGCCGAGATTCATCCCATCTT

63 MetValSerGluAlaAlaArgPheThrPheIleGlu

03 ATACATTACAGCGGAGTCATTCTTCCGACA

```

179 n.....AnphegIngnIargIleargProIaAsnaSnrHlleSerL 194
242 TGAGACTGATTGGGCGCACATCCAGCAATCCAGCACTCT...ACG 288
194 euIuInLysrIpgIyLysLeuSerPhegInIleargThrSerGlyAla 210
289 GATGCGTTTAAATACCATTCCTGGTGTATATCCAGCGTAACCTT 338
211 AsnGlyMetPheSerGluAlaValGluLeuGluAlaValAsnGlyLysI 227
339 CGTACGATGACGATGTCGCGAGTATGATGCGAGCTTACGATCAGT 388
227 sTyTyTyValThrAlaValAspGlnValIySproIyIleAlaLeuL 244
389 TGTTGTATGTAGGAC 405
244 yspheValcysLysasp 249

```

seq_name: /cgn2_6/ptodata/1/paa/US09_NFM_COMB.pep:us-09-663-826-7

seq_documentation_block:

```

; Sequence 7, Application US/09663826
; GENERAL INFORMATION:
; APPLICANT: Russel, David R.
; TITLE OF INVENTION: METHOD FOR PRODUCING ANTIBODIES AND
; CORRESPONDENCE ADDRESSES:
; NUMBER OF SEQUENCES: 15
; ADDRESS: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/663,826
; FILING DATE: 15-Sep-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/902,486
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 670513.90261
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 7:
; LENGTH: 248 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-663-826-7

```

alignment_scores:

```

Quality: 145.50 Length: 135
Ratio: 1.565 Gaps: 5
Percent Similarity: 68.889 Percent Identity: 31.111

```

alignment_block:

US-09-627-165b-15 x US-09-663-826-7 ..

Align seq 1/1 to: US-09-663-826-7 from: 1 to: 248

```

1 TACACAGATCTGGAGCGATACGCCGGTCAAT...AGGAGACAGATCCCTCT 47
112 TYGUAATGLeuGlnThrAlaAlaGlyLysIleArgGluAsnIleProLe 128
48 GGGTAAAGAGAACTCAATTCATCCGCTCGGCGCTTGTATCCAGGCG 97
128 uGlyLeuProAlaLeuAspSerAlaIleThrThrLeuTyTyTyThrAla 145
98 GCAGACCCCGGCGCCAGCTGCTCCCTTAATCTCTTACATGATGATC 147
145 laSerSerAlaIaSerAla.....LeuValLeuIleIleSerThr 159
148 TCCGAGCGCCGCAATTCATTCCTTTGGAGGCGTCGCCAATFACAT 197
160 AlaGluSerAlaArgTyTyLysPheIleGluGlnIleGlyLysArgVa 176
198 TACACCGGGGAGTCATTTCTCCGACATGTACATGCTGAGCTGAGAGA 247
176 laSp.....LysThrPheLeuProSerIleuAlaThrIleSerLeuGluA 191
248 CTAGTGGGCGCCACATCCAGCAATCCAG.....CAGTCTACGAT 291
191 snSotrPserAlaLeuSerLysGlnIleGlnIleAlaSerThrAsn 207
292 GCGCTTTTAAATACCATTCCTGGTGTATATCCAGCGTAACCTTGT 341
208 GlyGlnPheGluSerProValIleuIleAspGlyAsnAsnGlnArgVa 224
342 GAGCTTGACANT.....GTTCGCGAGTATCGGACGACCTTAGCATCA 385
224 lSerIleThrAsnAlaSerAlaArgValIleThrSerAlaIleAlaLeuL 241
386 TGTTG 390
241 leuL 242

```

seq_name: /cgn2_6/ptodata/1/paa/US09_NFM_COMB.pep:us-09-663-826-15

seq_documentation_block:

```

; Sequence 15, Application US/09663826
; GENERAL INFORMATION:
; APPLICANT: Russel, David R.
; TITLE OF INVENTION: METHOD FOR PRODUCING ANTIBODIES AND
; CORRESPONDENCE ADDRESSES:
; NUMBER OF SEQUENCES: 15
; ADDRESS: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/663,826
; FILING DATE: 15-Sep-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/902,486
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 670513.90261
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000

```


TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 496 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 15:
US-09-663-826-15

alignment_scores:
Quality: 145.50 Length: 135
Ratio: 1.565 Gaps: 5
Percent Similarity: 68.889 Percent Identity: 31.111

alignment_block:
US-09-627-165B-15 x US-09-663-826-15 ..

Align seg 1/1 to: US-09-663-826-15 from: 1 to: 496

```
1 TACACAGATCTGAGCGATACCGCGTCAT...AGGACCAAGATCCCTCT 47
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
112 TyrgluarglenglnthrAlaAlaGlySilearglusnIleProle 128
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
48 GGGTATAGAGAACTCAATCCGTCGCGCTGCTTATCCAGGCG 97
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
128 uGlyleuProAlaLeuaspSerAlaIlethrThleuYtyrThra 145
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
98 GCAGACCCCGGGCCCAAGCTGTCCTTATATCCATTCAAGATGATC 147
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
145 laseSerAlaIaIaSerAla.....leuValIleuIleGlnSerThr 159
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
148 TCCGAGCCCGGAGATTCATCCATCTTTGGAGGCGTCCCAATACAT 197
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
160 AlagluSerAlaArgYtyrLysPheIleGluGlnIleGlyLysArgVa 176
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
198 TAAACGCGGGAGTCATTTCCCGCATGTATACGTCGAGCGAGAGA 247
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
176 lasp.....LysThrPheLeuProSerleuAlaThrIleSerleuGlnVa 191
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
248 CTAGTTGGGGCCCAACATCCACGCAATGCAG.....CACTTACGAGAT 291
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
191 snAsnTrpSerAlaLeuSerLysGlnIleGlnIleAlaSerThrAsnSn 207
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
292 GCGCGTTTAAATACCACTTCGTTGGGTATATCCACGGTAACCTGCT 341
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
208 GlyGlnPheGluSerProValValIleuIleaspGlyAsnAsnGlnArgVa 224
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
342 GACGTTGAGCAAT.....GTTCCGAGCGTGTATCCGACGCTTACGATCA 385
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
224 lberIlethrAsnAlaSerAlaArgValValThrSerAsnIleAlaLeuL 241
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
386 TCTTG 390
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
241 euleu 242
```

seq_name: /cgn2_6/ptodata/1/paa/us10_NEW_COMB.pep:US-10-127-890-4

seq_documentation_block:
; Sequence 4, Application US/10127890
; GENERAL INFORMATION:
; APPLICANT: Better, Marc D.
; Carolli, Stephen F.
; Studnik, Gary M.
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
; Proteins
; NUMBER OF SEQUENCES: 173
; CORRESPONDENCE ADDRESS:
; ADDRESS: McAndrews, Held & Malloy, Ltd.
; STREET: 500 West Madison Street, 34th floor
; CITY: Chicago
; STATE: Illinois

COUNTRY: USA
ZIP: 60661
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/127,890
FILING DATE: 23-Apr-2002
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/646,360
FILING DATE: 13-MAY-1996
APPLICATION NUMBER: PCT/US94/05348
FILING DATE: 12-MAY-1994
APPLICATION NUMBER: US 08/064,691
FILING DATE: 12-MAY-1993
APPLICATION NUMBER: US 07/988,430
FILING DATE: 09-DEC-1992
APPLICATION NUMBER: US 07/901,707
FILING DATE: 19-JUN-1992
APPLICATION NUMBER: US 07/787,567
FILING DATE: 04-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: McNicholas, Janet M.
REGISTRATION NUMBER: 32,918
REFERENCE/DOCKET NUMBER: 200-70.P4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/707-8869
TELEFAX: 312/707-9155
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 263 amino acids
MOLECULE TYPE: protein
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-127-890-4

alignment_scores:
Quality: 141.00 Length: 148
Ratio: 1.454 Gaps: 6
Percent Similarity: 65.541 Percent Identity: 30.405

alignment_block:
US-09-627-165B-15 x US-10-127-890-4 ..

Align seg 1/1 to: US-10-127-890-4 from: 1 to: 263

```
1 TACACAGATCTGAGCGATACCGCGTCAT...AGGACCAAGATCCC 44
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
109 TyrgluasnleuGlnThr..AlaAlaHisLysIleArgGlnsnIleAs 124
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
45 TCTGGTATAGAGAACTCAATCCGTCGCGCTTGTATATCCAG 94
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
124 PleuGlyleuProAlaLeuSerSerAlaIlethrThrleuPheYtyrVa 141
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
95 GCGGACGACCCCGGGCCCAAGCTGTCCTTATATATCCATTCACATG 144
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
141 snAlaGlnSerAlaProSerAla.....leuValIleuIleGlnThr 155
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
145 ATCTCCGAGCGCGGAGATTCATCCATCTTTGGAGGCGTCGCCAATA 194
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
156 ThrAlaGlnAlaAlaArgPheYtyrIleGluArgHisValAlaLysTy 172
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
195 CATTAACAGCGGGAGATCTTCTCCGACATGTATACGTCGAGCTGC 244
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
172 Val.....AlaThrAsnPheLysProAsnleuAlaIleIleSerleuG 187
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
245 AGACTAGTTGGGGCCCAACATCCACGCAATGC.....CAGCAATCTACG 288
```

```
187 ||:::||||: || ||::: ||:::
    luAsnGlnItrpSerAlaLeuSerLysGlnIlePheLeuAlaGlnAsnGln 203
289 GATGGCGTTTAAATACCATTTGTTGGGTATATCCACCGTAACTT 338
    || ||:::||||| || ||:::|||||
204 GlyGlyLysPheArgAsnProValAspLeuIleLysProThrGlyGluAr 220
    ::::||||| :::: ||::: ||
339 CGTGACGTTGACCAATGTTCCGCGACGTGATCGCAGCTTACGATCARGT 388
    ::::||||| :::: ||::: ||
220 gPheGlnValThrAsnValAspSerAspValValLysGlyAsnIleLysL 237
389 TGTGTATGTAGGACGACCATCTTCTCCGACGTCCCTAT 432
    ||::: ||::: ||::: ||::: ||::: ||:::
237 euLeuLeu...AsnSerArgAlaSerThrAlaAspGluAsnPhe 250
```